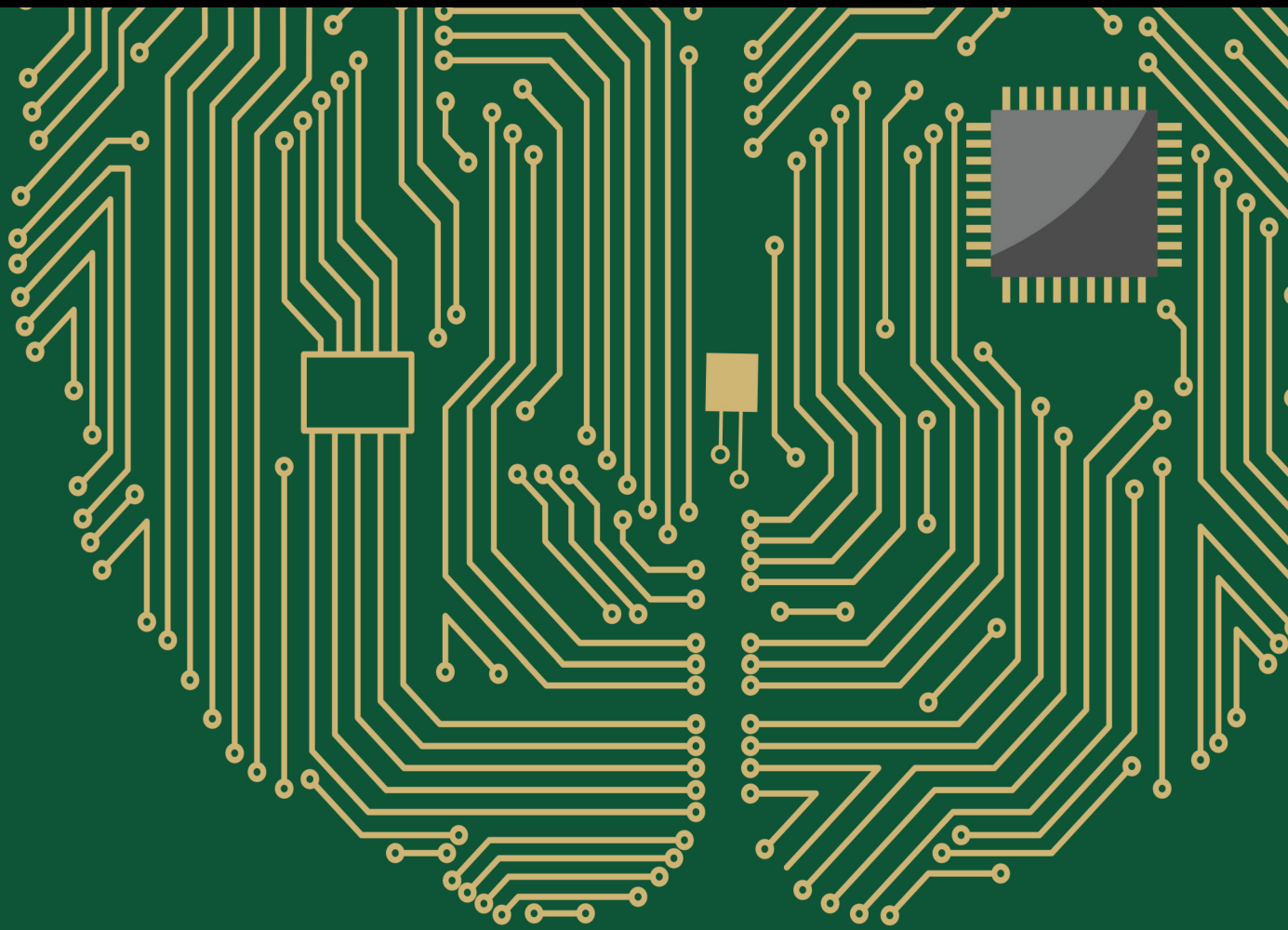# Recent Advances in Multimodal Environment for Biomedical Diagnosis and Computational Analysis

Lead Guest Editor: Amandeep Kaur
Guest Editors: Shivani Goel and Wattana Viriyasitavat

Recent Advances in Multimodal Environment for Biomedical Diagnosis and Computational Analysis

Recent Advances in Multimodal Environment for Biomedical Diagnosis and Computational Analysis

Lead Guest Editor: Amandeep Kaur
Guest Editors: Shivani Goel and Wattana Viriyasitavat

Contents

Contents

Contents

Contents

Retraction

Retracted: Analysis of Signs and Effects of Surgical Breast Cancer Patients Based on Big Data 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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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] Z. Hong, Q. Xu, X. Yan, R. Zhang, Y. Ren, and Q. Tong, "Analysis of Signs and Effects of Surgical Breast Cancer Patients Based on Big Data Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3373553, 8 pages, 2022.

Retraction

Retracted: Behavior Data Analysis of English Learners Based on Discrete Dynamic System Modeling 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] H. Chen and Y. Gao, "Behavior Data Analysis of English Learners Based on Discrete Dynamic System Modeling Method," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5409571, 8 pages, 2022.

Retraction

Retracted: Visualization Method of Key Knowledge Points of Nursing Teaching Management System Based on SOM Algorithm and Biomedical Diagnosis

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 L. Fang, "Visualization Method of Key Knowledge Points of Nursing Teaching Management System Based on SOM Algorithm and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7057437, 9 pages, 2022.

Retraction

Retracted: The Impact of Education Based on New Internet Media Technology on College Students' Mental Health and Biomedical Diagnosis

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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Wang, W. Wei, and J. Zhao, "The Impact of Education Based on New Internet Media Technology on College Students' Mental Health and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3617938, 7 pages, 2022.

Retraction

Retracted: Mobility and Quality of Life in Chinese Elderly and Geriatric Patients and Biomedical Diagnosis

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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Xin and L. Li, "Mobility and Quality of Life in Chinese Elderly and Geriatric Patients and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7457161, 9 pages, 2022.

Retraction

Retracted: Analysis of Structured Data in Biomedicine Using Soft Computing Techniques and Computational Analysis

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. H. Rahman, "Analysis of Structured Data in Biomedicine Using Soft Computing Techniques and Computational Analysis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4711244, 11 pages, 2022.

Retraction

Retracted: Detection of Heart Arrhythmia on Electrocardiogram using Artificial Neural Networks

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] M. Badr, S. Al-Otaibi, N. Alturki, and T. Abir, "Detection of Heart Arrhythmia on Electrocardiogram using Artificial Neural Networks," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1094830, 10 pages, 2022.

Retraction

Retracted: Assessment of Ecological Environment Quality for Urban Sustainable Development Based on AHP

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. Zhou, "Assessment of Ecological Environment Quality for Urban Sustainable Development Based on AHP," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4056713, 9 pages, 2022.

Retraction

Retracted: A Plain Bayesian Algorithm-Based Method for Predicting the Mental Health Status and Biomedical Diagnosis of University Students

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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Wang, "A Plain Bayesian Algorithm-Based Method for Predicting the Mental Health Status and Biomedical Diagnosis of University Students," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2617488, 5 pages, 2022.

Retraction

Retracted: Neural Network Based on Health Monitoring Electrical Equipment Fault and Biomedical Diagnosis

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. Zhang and Y. Lyu, "Neural Network Based on Health Monitoring Electrical Equipment Fault and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8358794, 7 pages, 2022.

Retraction

Retracted: The Influence of Martial Arts on Spine CT Image Morphological Structure Based on Optimized Ant Colony 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Yuhuan, H. Xueting, and W. Weiyue, "The Influence of Martial Arts on Spine CT Image Morphological Structure Based on Optimized Ant Colony Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2725819, 10 pages, 2022.

Hindawi

Retraction

Retracted: Smart Health Monitoring System with Wireless Networks to Detect Kidney Diseases

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. Dhanke, N. Rathee, M. S. Vinmathi, S. Janu Priya, S. Abidin, and M. Tesfamariam, "Smart Health Monitoring System with Wireless Networks to Detect Kidney Diseases," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3564482, 11 pages, 2022.

Research Article

A Greedy Optimized Intelligent Framework for Early Detection of Alzheimer's Disease Using EEG Signal

R. Swarnalatha [ID]

Department of Electrical & Electronics Engineering, Birla Institute of Technology & Science, Pilani, Dubai Campus, Dubai, UAE

Correspondence should be addressed to R. Swarnalatha; swarnalatha@dubai.bits-pilani.ac.in

Recent researchers have been drawn to the analysis of electroencephalogram (EEG) signals in order to confirm the disease and severity range by viewing the EEG signal which has complicated the dataset. The conventional models such as machine learning, classifiers, and other mathematical models achieved the lowest classification score. The current study proposes to implement a novel deep feature with the best solution for EEG signal analysis and severity specification. A greedy sandpiper-based recurrent neural system (SbRNS) model for predicting Alzheimer's disease (AD) severity has been proposed. The filtered data are used as input for the feature analysis and the severity range is divided into three classes: low, medium, and high. The designed approach was then implemented in the matrix laboratory (MATLAB) system, and the effectiveness score was calculated using key metrics such as precision, recall, specificity, accuracy, and misclassification score. The validation results show that the proposed scheme achieved the best classification outcome.

1. Introduction

AD is a neurological malfunction identified by debilitation of analytic functions and amnesia [1]. This is mainly due to the unusual increase of protein around brain cells [2]. AD is often considered as a usual aging process and affects older people more than younger ones [3]. In humans, AD signs are increased confusion and loss of learning ability and memory [4]. Generally, AD is categorized into three stages based on its symptoms and effects [5]. In the beginning stage (mild AD), the sign of AD is most commonly amnesia [6], which does not change a person's daily life. The next phase is considered moderate AD [7], which is recognized by increased confusion and difficulty in learning [8]. In this phase, patient's dependence on the other people increases [9]. The third stage is severe AD, which is characterized by the entire debilitation of individual [10]. The early identification of AD is important in the case of mild AD and moderate AD [11, 12]. Detecting this disease is quite difficult because signs like amnesia are often considered as usual aging signs [13]. Moreover, the imaging and signal analysis system has been introduced in the digital field to analyze the

AD severity range with the help of artificial intelligence (AI) [14, 15]. Hence, several neural models and classifiers were implemented to find the AD severity rate in the early stages without the requirement of high resource usage [16].

Generally, disease detection is done by large testing and eliminating the other disease cases [17]. Neurological testing, physiological and psychological examination, and blood test help to detect AD [18]. The basic Alzheimer's disease detection with AI is represented in Figure 1. Perform brain scans, such as computed tomography (CT), magnetic resonance imaging (MRI), or positron emission tomography (PET), to support an Alzheimer's disease diagnosis or rule out other possible causes for symptoms. Psychological testing such as mini-mental state testing and image techniques based on AI were developed to detect AD [19]. Still, they faced several demerits such as inaccurate disease classification, high error rate, and so on [20]. Thus, the researchers paid attention to AD detection techniques based on the analysis of EEG signals [21]. Here, the brain's reading is recorded using electrodes directly by 10–20 electrode systems [22]. AD diagnosis based on EEG signals provides early detection and accurate disease classification [23].
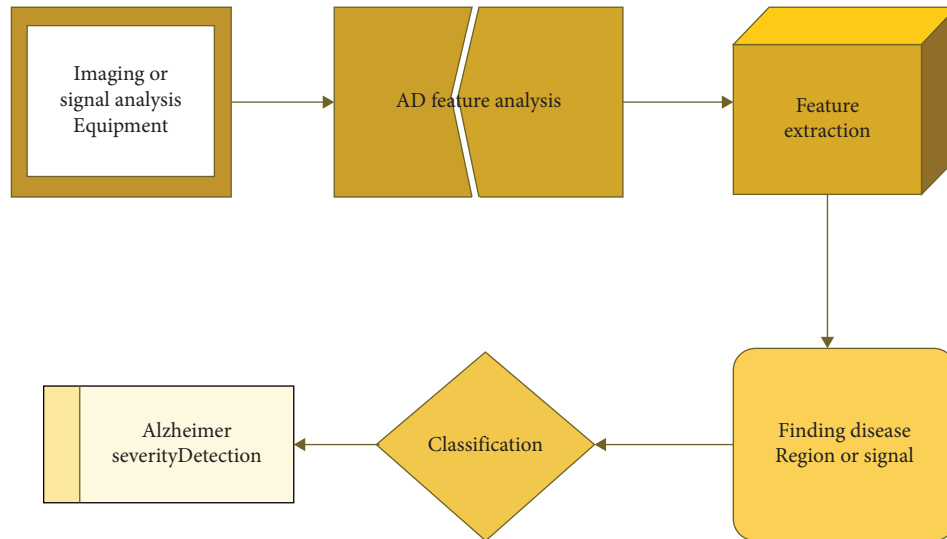
Figure 1: Basic Alzheimer detection using AI.

Recently, many researchers tried to detect AD in the early stage by identifying changes in the synchrony of EEG signals [24]. However, the diagnosis is not accurate in the case of single synchrony [25]. In addition, for the EEG signal-based AD severity analysis, different and several brain samples are taken from one person [26]. This has helped to find the brain neuron movements based on their actions and emotions. The EEG-based AD detection technique is used as a tool to diagnose AD early for a huge number of people [27].

To overcome the demerits faced by these techniques, a highly efficient and accurate AD detection was presented in this paper using EEG signals [28, 29]. In addition, a novel deep network has been executed by incorporating the optimal solution for detecting disease range. Also, in the designed model, the noise removal function is executed in the hidden phases that have afforded the finest noise feature tracking and neglecting. The present noises were eliminated properly in the filtering stage, so the algorithm complexity was minimized, and the classification score was maximized. Finally, a comparison assessment was conducted to check the improvement score in analyzing the AD severity range.

The current research is organized as follows. Section 2 describes the existing works in detail. The problem related to analyzing the AD EEG signal is defined in Section 3. The proposed scheme and the solution of the problem are mentioned in Section 4. The proposed model's outcome is discussed in Section 5, and research arguments are concluded in Section 6.

2. Related Works

Some of the recent works related to AD detection are described below.

Safi and Safi [27] developed a technique by using Hjorth parameters to diagnose AD in its early stage. This model uses different filtering techniques such as empirical mode decomposition and discrete wavelet transform to filter the errors in the brain signals. Also, the K-nearest neighbor algorithm was applied to classify the diseases. In past decades, different neuropsychological testing based on magnetic resonance imaging (MRI) was developed to diagnose AD. Thus, Murugan et al. [28] developed a MRI-based convolutional neural network model to classify and detect AD accurately. This model is highly efficient in detection and classification of diseases. But, they did not provide accurate AD detection in case of a large population.

The interface between the features of EEG signal and AD is not clarified in existing detection models. Therefore, Li et al. [29] presented a novel technique to establish the relationship between AD detection and EEG signal which presents latent parameters in 3D phase and also provides high classification efficiency. However, the computational cost is high in this model. Deepa and Chokkalingam [30] presented an optimized VGG-16 framework to classify AD using arithmetic optimization which uses MRI-based images for disease classification. Here, the AD was classified as mild, moderate, and severe. Moreover, the computational complexity and cost are reduced in this technique. However, the experimental outcome of this technique for a large population does not provide efficient classification.

Recently, EEG signals have been widely used in medical fields to diagnose the disease in its early phase. Hence, Cejnek et al. [31] designed a model to detect mild AD based on EEG records. This technique uses an adaptive filter to preprocess the recorded EEG dataset. In predictive function, the linear NN technique with gradient adaption is used. This technique is validated by testing the data of AD patients. However, it does not attain high sensitivity and accuracy.

3. AD System Model and Problem Description

Usually, the disease forecasting process using signaling data is more complicated because of the feature variations in the minute level. Hence, the normal classification and signal analysis model has required more resources to find the feature variation and to identify the disease signal. In many
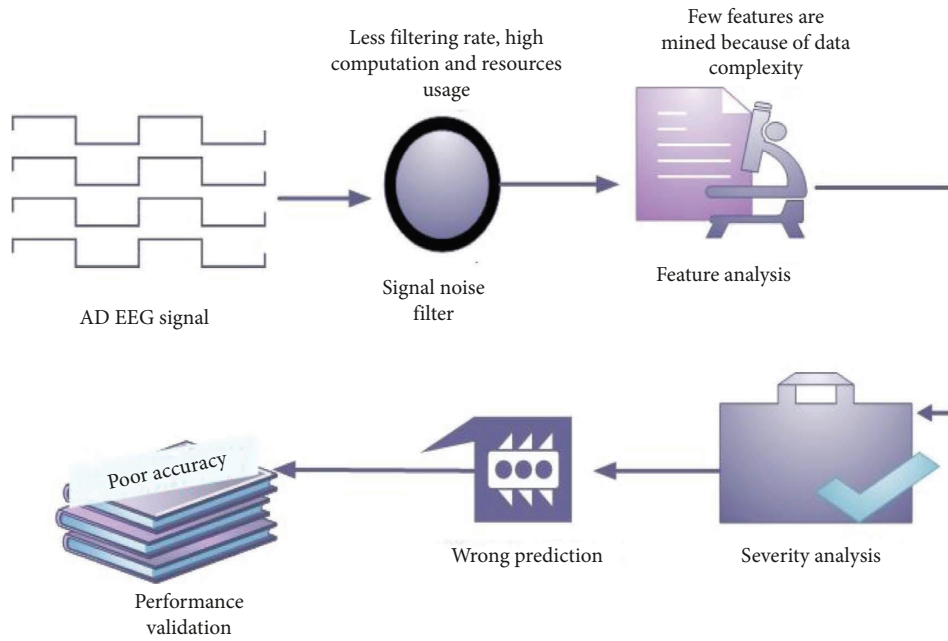
Figure 2: AD system model with problem.
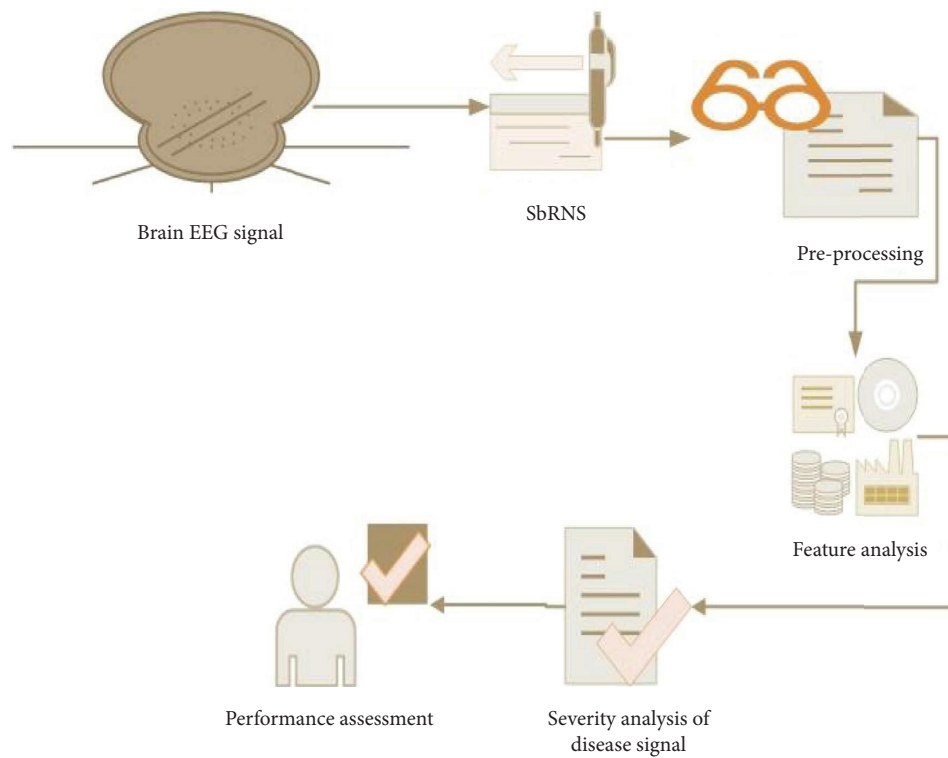
Figure 3: Proposed SbRNS architecture.

cases, the algorithm complexity has been recorded because of signal analysis delay. These issues have motivated this work toward presenting an optimized deep neural system for EEG signal analysis and disease signal feature classification. The disease that was considered for validating the developed model is AD [32].

Analyzing the brain signal is the foremost topic in the medical industry to detect and treat different kinds of diseases. Especially, analyzing the EEG signal and identifying the disease features and their severity range are highly complicated because of the noisy unbalanced data. The problems analyzing in EEG signal are described in Figure 2.
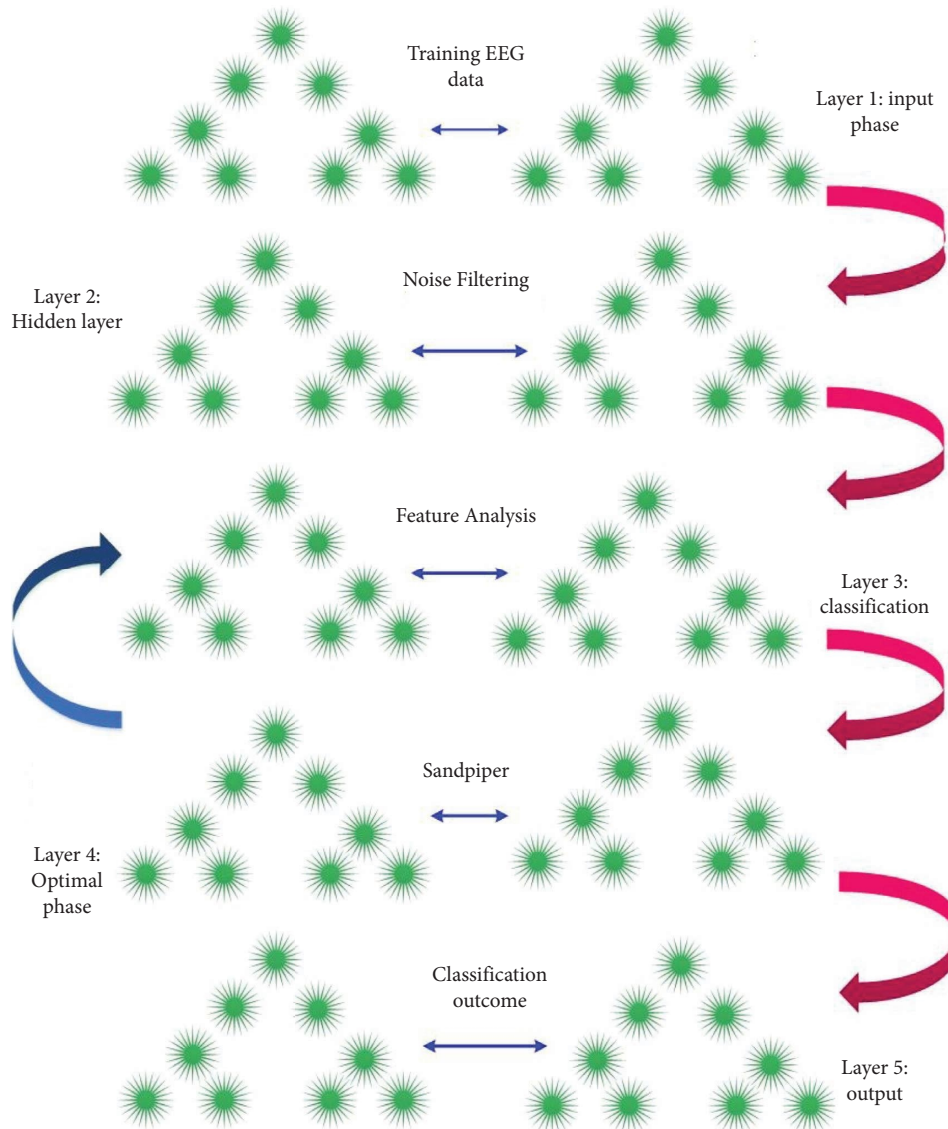
FIGURE 4: Layers of SbRNS.

A novel sandpiper-based recurrent neural system (SbRNS) has been introduced for predicting Alzheimer's disease in an earlier stage using EEG signal. Primarily, the EEG signals were collected and trained to the system and then the preprocessing function was activated to filter the present noise in the trained datasets. Consequently, the feature analysis was performed, and the disease features were extracted and then the severity range was measured. Finally, the key parameters were estimated and compared with the other schemes.

4. Proposed Methodology

The proposed architecture is illustrated in Figure 3. In the feature analysis steps, the required features were traced and mined based on the specification of 0th class and 1st class.

Here, the signal pixels were analyzed and traced for the 0th class features and then the extraction was done. Consequently, the disease specification and severity range estimation were performed. Finally, the comparison analysis helped to measure the performance enhancement by the developed scheme.

4.1. Proposed SbRNS Design. The planned model has five different phases that includes data importing, error filtering, classification module, optimal phase, and output phase. The data importing process is executed in the training phase of the initial layer. Noise elimination function is performed in the second phase hidden layer. Feature analysis and the severity level identification are processed in the third classification phase, and then the classification phase parameters are tuned by the optimal layer. In the fourth layer, the optimal phase sandpiper fitness is utilized. Finally, the severity forecasting outcome is recorded in the fifth layer, output phase.

The proposed scheme functions on the basis of the sandpiper approach [33] and recurrent neural scheme [34].

These layers are elaborated in Figure 4. Here, the reason for incorporating the sandpiper function in the recurrent neural classification phases is to earn the finest severity forecasting score.

4.2. Preprocessing Model.

The function preprocessing is designed to eliminate the training flaws from the trained sets. This process tends to gain the finest accurate severity analysis score and less computational complexity. Moreover, the data importing process has been processed by (1). Here, the AD brain EEG signal dataset is determined as b_e, dataset training function is determined as $F(b_e)$, and the n number of EEG signals is described as $1, 2, 3, 4, \ldots, n$:

$$F(b_e) = b_e\{1, 2, 3, 4, \cdots, n\}. \tag{1}$$

The data training function is performed in the input phase of the novel SbRNS. After training the data, the function preprocessing is activated in the hidden phase to eliminate the noise features.

$$J(b_e) = |b_e(q, a) - b_e(a)|, \tag{2}$$

where J represents the preprocessing variable, q represents the normal features in the database, and the noise features are described as a. The noise features that are present in the trained database are removed by the preprocessing process that is represented as $b_e(q, a) - b_e(a)$.

4.3. Feature Analysis.

The output of the preprocessing layer is earned as the error-free data. Then, those data are considered as the input of the feature analysis process. Forecasting the disease or severity range in the entire data is difficult and has maximized the computational complexity. Considering this problem, the process feature analysis has been executed to track and extract the present features in the trained sets.

$$F^*(b_e) = Z - \left(\frac{y(b_e)}{\max}\right), \tag{3}$$

where the AD signal features are described as Z and F^* determines the feature analysis variable. The process of feature extraction is indicated in (3). Here, Z represents the 0th class. In this, the 0th class features are traced and mined. Also, the unwanted features y are neglected because y falls under the 1st class. In (3), max denotes the maximum possible iterations. This feature analysis process is executed by the fitness of sandpiper.

4.4. Classification Module.

After the feature analysis function, the severity range of the AD has been predicted in the form of low, high, and medium. The analyzed features Z are categorized under the 0th class and then AD range is specified as high severity; if Z falls under the 1st class, then it is medium severity. If Z is not falling under classes 0 and 1, then it is specified as low severity.
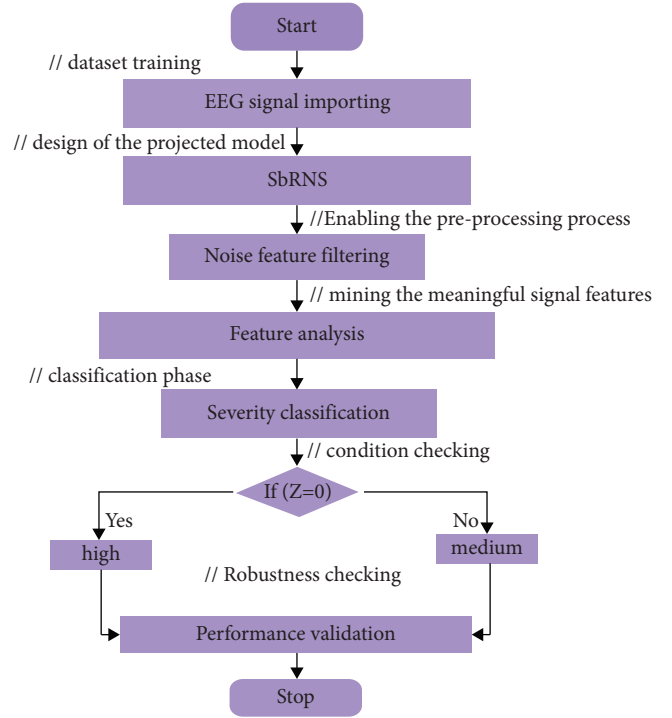
Figure 5: Flowchart of proposed SbRNS.

$$D^*(b_e) = \begin{cases} if\ (Z = 0), & \text{high,} \\ if\ (Z = 1), & \text{medium,} \\ \text{else,} & \text{low,} \end{cases} \tag{4}$$

where D^* represents the classification parameter. Hence, based on the "if" condition, the severity of the AD is specified as high and medium. The classification process is valued in (4).

The defined formulation of the novel SbRNS is represented in Algorithm 1. By processing each step of SbRNS algorithm, the EEG signal was analyzed, and the severity score was forecasted. In addition, the flow of the steps in executing the present scheme is elaborated in Figure 5.

5. Results and Discussion

The discussed model is executed in the MATLAB programming platform and processed in Windows 10 platform. Initially, the EEG signal of AD is gathered from the standard site and taken as the input of the MATLAB system. The parameter specifications are indicated in Table 1.

In addition, the EEG signal processing framework functions in the ratio of $75 : 25$, i.e., training 75% and testing 25%. The class that has been considered in the severity specification is low, medium, and high.

The ability of the model in specifying the severity range is described through the validation, and the graphical representation is given in Figure 6. The validation graph has reached 100% accuracy range that described the exact prediction, and the loss is very low.

To measure the working range of the proposed design, a few different test samples are considered that are normal

```
start
{
    int b_e = 1, 2, 3, . . . , n;
    //initializing the EEG signal data
    Preprocessing ()
    {
        int j, q, a;
        //initialing the filtering parameters
        j ⟶ remove_a (b_e)
        //Eliminating the noise features
    }
    Feature analysis ()
    {
        int Z, y;
        //activating the feature analysis variables
        F* ⟶ Z(b_e)
        //meaningful signal features were traced and extracted
    }
    Severity Classification ()
    {
        f (Z = 0)
        {
        High severity
        }else if (Z = 1)
        {
        medium
        }else (low)
        //hence, the severity of the AD has been measured
    }
}
Stop
```

ALGORITHM 1: SbRNS.

signal, low AD signal, medium AD signal, and high AD signal. Here, the normal signal is also considered for this experiment to check the working rate of the presented design. The results of the forecasted normal brain signal are shown in Figure 7.

Different EEG signal samples were tested. The outcomes of medium AD, high AD, and low AD are indicated in Figures 8–10.

5.1. Performance Validation.

Recall validation is an evaluation of the sensitivity range in classifying the severity categories. Also, it has provided the stability range of the executed model in classifying the disease range in the presence of false prediction and actual prediction. Thus, if a system has recorded a good recall measures, it is better in severity specification. The recall is was formulated as

$$\text{recall} = \frac{\text{true positive}}{\text{false negative} + \text{true positive}}. \quad (5)$$

To measure the positive values in the severity prediction, the precision metrics were valued. Hence, this has afforded the actual forecasting outcome in severity range specification. The precision parameter is defined as

$$\text{Precision} = \frac{\text{true positive}}{\text{false positive} + \text{true positive}}. \quad (6)$$

The exactness of severity specification is determined as accuracy. Hence, the gained accuracy has verified the robustness of the designed approach in specifying the Alzheimer severity range. The accuracy is defined as

$$\text{accuracy} = \frac{\text{exact forecast}}{\text{total forecast}}. \quad (7)$$

The mean efficiency of the severity classification process has been found in the F-measure validation. The summation and product of the recall and precision have described the F-score, which is indicated as

$$F_{\text{Value}} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}. \quad (8)$$

The metrics specification was estimated to assess the negative scores in predicting the AD severity range as follows:

$$\text{specificity} = \frac{\text{true negative}}{\text{false positive} + \text{true negative}}. \quad (9)$$

Here, the specificity score was determined by only taking into account negative classifications, i.e., a true negative and
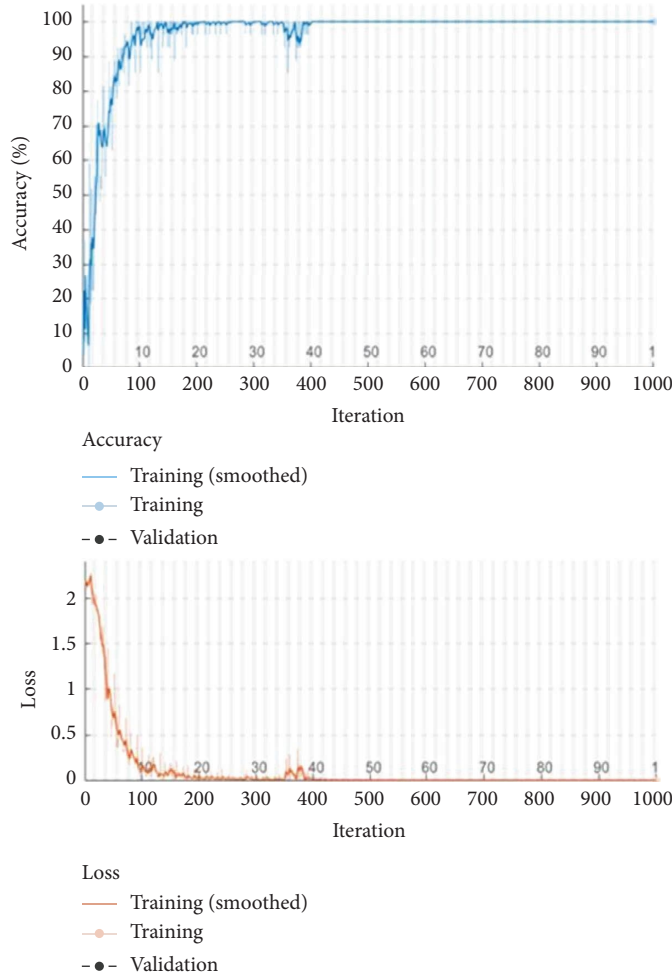
Results
Validation accuracy: N/A
Training finished: Reached final iteration

Training Time
Start time: 10-Jun-2022 16: 44: 57
Elapsed time: 2 min 34 sec

Training Cycle
Epoch: 100 of 100
Iteration: 1000 of 1000
Iterations per epoch: 10
Maximum iterations: 1000

Validation
Frequency: N/A

Other Information

Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.001

Accuracy

— Training (smoothed)
—•— Training
–•– Validation

Loss

— Training (smoothed)
—•— Training
–•– Validation

Figure 6: Validation of loss and accuracy.
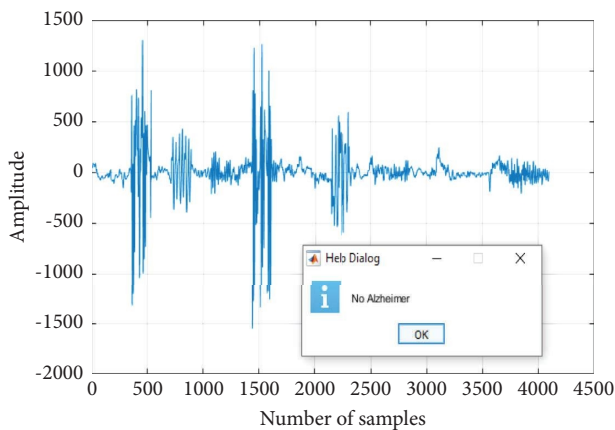
Figure 7: Normal EEG signal.
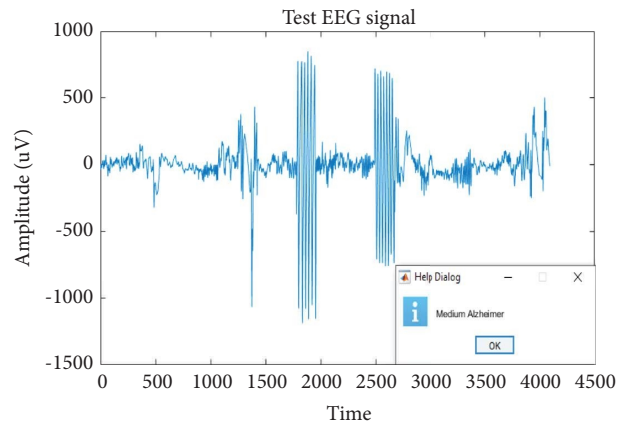
Figure 8: Classified medium severity AD.

a false positive. Hence, the overall validation of the developed novel SbRNS was described in dual phases that are with and without incorporating the sandpiper fitness solution in the recurrent model, which is graphically defined in Figure 11. Before applying the sandpiper model in the recurrent system, the recorded specificity score was 98%; after adding

Table 1: Details of execution parameters.

Parameter specifications	
Programming language	MATLAB
Version	R 2020 b
Operating system	Windows 10
Dataset type	AD signal
Signal type	EEG
Total EEG signals	11500
Objective	Severity range prediction

TABLE 2: Comparison assessment.

| | Comparison statistics | | | |
	Precision (%)	Specificity (%)	F-measure (%)	Accuracy (%)
EIM	99.69	99.6	98.9	98.7
KNM	96.7	95.7	98.3	98.1
SbRNS (B.O)	99.6	99.3	99	99
SbRNS (A.O)	99.8	99.9	99.5	99.8
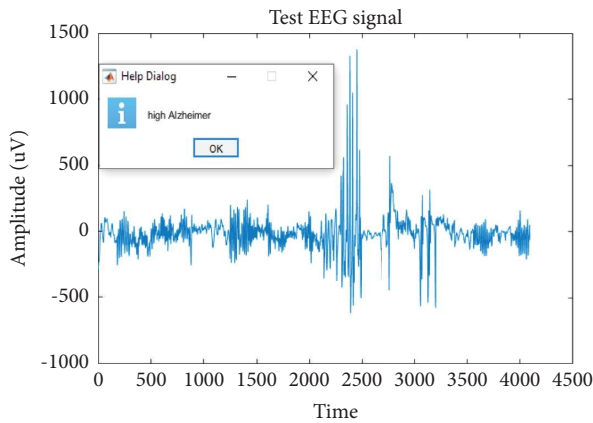
FIGURE 9: Classified high severity AD.
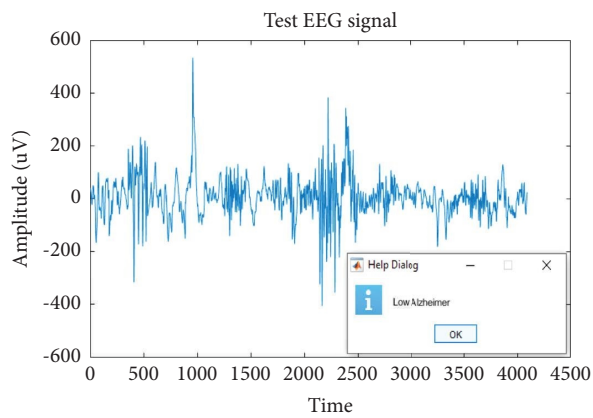
FIGURE 10: Classified low severity AD.

FIGURE 11: Performance assessment.
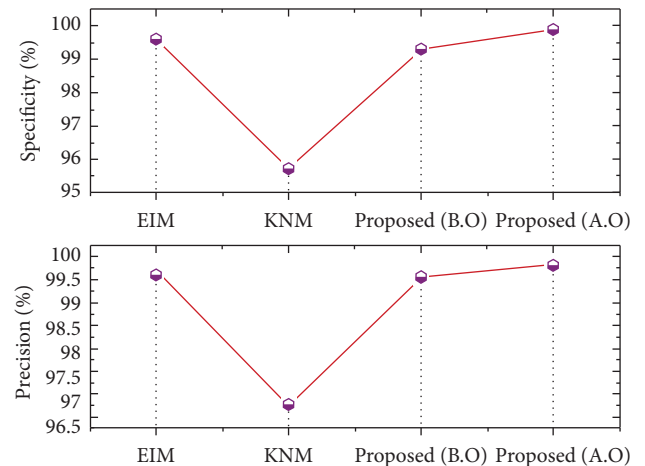
FIGURE 12: Assessment of specificity and precision.

the sandpiper function in the classification phase of the recurrent system, 99.5% of recall was recorded.

5.2. Comparison Assessment.

The conventional schemes that have been taken used to validate the designed model are the extreme intelligent model (EIM) and K-nearest model (KNM) [35]. Besides, the presented approach is measured in dual phases that are before and after applying the optimization model.

The EIM has attained a precision score as of 99.69% and specificity of 99.6%; the approach KNM has earned a specificity rate as of 95.7% and precision of 96.7%. Moreover, before incorporating the sandpiper fitness, the presented scheme SbRNS has recorded the specificity score as of 99.3%

and precision of 99.6%. After applying the fitness solution of sandpiper function model, 99.8% precision and 99.9% specificity were reported.

The F-measure and accuracy rate gained by the scheme EIM are 98.9% and 98.7%, respectively; the approach KNM has gained an accuracy score of 98.1% and F-score of 98.3%. In addition, before performing the sandpiper function, the recorded F-score is 99% and accuracy is 99%. Furthermore, after executing the sandpiper model, the recorded F-score is
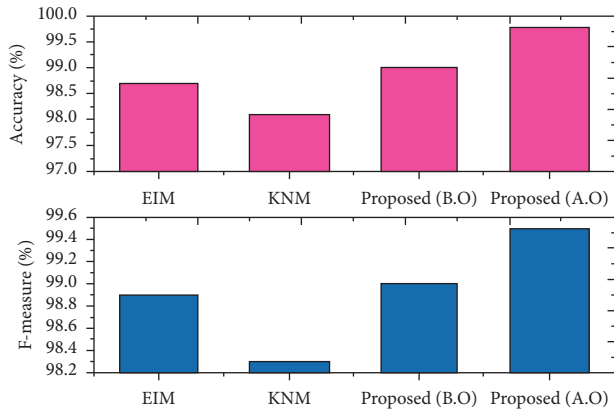
FIGURE 13: Assessment of accuracy and *F*-measure.

TABLE 3: Performance assessment of SbRNS.

Overall validation	
Precision (%)	99.8
Specificity (%)	99.9
F-score (%)	99.5
Accuracy (%)	99.8
Recall (%)	99.95
Error (%)	0.1

99.5%, and the exactness score is 99.8%. The statistics before and after the optimization are shown in Figure 12.

By comparing the recent scheme, the present approach has earned the best outcome in classifying the disease severity range. This has verified the effectiveness of the developed model in EEG signal analysis. In addition, after incorporating the sandpiper, the highest performance was recorded. It is clearly seen that the need of the optimal model in severity specification is highly recommended. The assessment of accuracy and *F*-measure is revealed in Figure 13. Also, the overall comparison assessment is described in Table 2.

5.3. Discussion. The validated parameters have described the proficient score of the designed model. The working performance of the proposed model is described by estimating all classification metrics that are precision, specificity, *F*-score, accuracy, recall, and error validation that are detailed in Table 3.

In this case, the metrics error was examined in order to examine the falling score of the designed approach. As a result, the earned misclassification score is only 0.1% percent, which is a very low and negligible state. This described the efficacy of the devised scheme.

6. Conclusion

A novel SbRNS has been implemented in the MATLAB framework for the classification of AD severity ranges using EEG signals, and the severity analysis procedure was carried out. The current SbRNS has a maximum severity forecast range of 99.8% when compared to other conventional approaches, and the classification rate was increased by 1%. The recorded sensitivity rate is 99.9%. In comparison to previous approaches, the specificity score has increased by 0.6% in comparison to earlier methods. The novel SbRNS was able to attain an extremely low error rate of 0.1%. Furthermore, when compared to the conventional recurrent neural model before applying the sandpiper function, the present SbRNS improved performance by 1% to 2%. As a result, the designed method is the a necessary scheme for EEG signal analysis and severity range prediction; however, it is not applicable to other application datasets. Creating a tuned hyper-parameter model will provide the best classification and feature analysis results in the future.

Data Availability

The data used to support the findings of this study are available online.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

References

[1] C. N. Metz and V. A. Pavlov, "Treating disorders across the lifespan by modulating cholinergic signaling with galantamine," *Journal of Neurochemistry*, vol. 158, no. 6, pp. 1359–1380, 2021.

[2] M. Martínez-Negro, G. González-Rubio, E. Aicart, K. Landfester, A. Guerrero-Martínez, and E. Junquera, "Insights into colloidal nanoparticle-protein corona interactions for nanomedicine applications," *Advances in Colloid and Interface Science*, vol. 289, p. 102366, 2021.

[3] A. Ferrari, D. Stark, F. A. Peccatori et al., "Adolescents and young adults (AYA) with cancer: a position paper from the AYA working group of the European society for medical oncology (ESMO) and the European society for paediatric oncology (SIOPE)," *ESMO open*, vol. 6, no. 2, p. 100096, 2021.

[4] T. T. Nguyen, T. T. D. Nguyen, T. K. O. Nguyen, T. K. Vo, and V. G. Vo, "Advances in developing therapeutic strategies for Alzheimer's disease," *Biomedicine & Pharmacotherapy*, vol. 139, p. 111623, 2021.

[5] S. L. Warren, A. A. Moustafa, and H. Alashwal, "Harnessing forgetfulness: can episodic-memory tests predict early Alzheimer's disease?" *Experimental Brain Research*, vol. 239, no. 9, pp. 2925–2937, 2021.

[6] E. S. Thoe, A. Fauzi, Y. Q. Tang, S. Chamyuang, and A. Y. Y. Chia, "A review on advances of treatment modalities for Alzheimer's disease," *Life Sciences*, vol. 276, p. 119129, 2021.

[7] P. T. Tseng, J. S. Jeng, B. S. Zeng et al., "Efficacy of non-invasive brain stimulation interventions in reducing smoking frequency in patients with nicotine dependence: a systematic review and network meta-analysis of randomized controlled trials," *Addiction*, vol. 76, p. 854, 2021.

[8] R. Knobler, P. Arenberger, A. Arun et al., "European dermatology forum: updated guidelines on the use of extracorporeal photopheresis 2020–Part 2," *Journal of the European*

Academy of Dermatology and Venereology, vol. 35, no. 1, pp. 27–49, 2021.

[9] J. V. Shanmugam, B. Duraisamy, B. C. Simon, and P. Bhaskaran, "Alzheimer's disease classification using pretrained deep networks," *Biomedical Signal Processing and Control*, vol. 71, p. 103217, 2022.

[10] H. Jayamohan, C. J. Lambert, H. J. Sant et al., "SARS-CoV-2 pandemic: a review of molecular diagnostic tools including sample collection and commercial response with associated advantages and limitations," *Analytical and Bioanalytical Chemistry*, vol. 413, no. 1, pp. 49–71, 2021.

[11] A. P. Porsteinsson, R. S. Isaacson, S. Knox, M. N. Sabbagh, and I. Rubino, "Diagnosis of early Alzheimer's disease: clinical practice in 2021," *The journal of prevention of Alzheimer's disease*, vol. 8, no. 3, pp. 371–386, 2021.

[12] M. Dashwood, G. Churchhouse, M. Young, and T. Kuruvilla, "Artificial intelligence as an aid to diagnosing dementia: an overview," *Progress in Neurology and Psychiatry*, vol. 25, no. 3, pp. 42–47, 2021.

[13] A. M. Alvi, S. Siuly, and H. Wang, "Neurological abnormality detection from electroencephalography data: a review," *Artificial Intelligence Review*, pp. 1–38, 2021.

[14] B. Llerena Zambrano, A. F. Renz, T. Ruff et al., "Soft electronics based on stretchable and conductive nanocomposites for biomedical applications," *Advanced Healthcare Materials*, vol. 10, no. 3, p. 20013, 2021.

[15] A. M. Alvi, S. Siuly, and H. Wang, "Neurological abnormality detection from electroencephalography data: a review," *Artificial Intelligence Review*, pp. 1–38, 2021.

[16] D. Puri, S. Nalbalwar, A. Nandgaonkar, and A. Wagh, "EEG-based diagnosis of Alzheimer's disease using Kolmogorov complexity," *Applied Information Processing Systems*, pp. 157–165, Springer, Singapore, 2022.

[17] V. Khurana, M. Gahalawat, P. Kumar et al., "A survey on neuromarketing using eeg signals," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 23, p. 897, 2021.

[18] P. M. Rodrigues, B. C. Bispo, C. Garrett, D. Alves, J. P. Teixeira, and D. Freitas, "Lacsogram: a new EEG tool to diagnose Alzheimer's disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3384–3395, 2021.

[19] X. Zhao, C. K. E. Ang, U. R. Acharya, and K. H. Cheong, "Application of Artificial Intelligence techniques for the detection of Alzheimer's disease using structural MRI images," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 2, pp. 456–473, 2021.

[20] Y. Qu, P. Wang, B. Liu et al., "AI4AD: artificial intelligence analysis for Alzheimer's disease classification based on a multisite DTI database," *Brain Disorders*, vol. 1, p. 100005, 2021.

[21] Z. Gao, W. Dang, X. Wang et al., "Complex networks and deep learning for EEG signal analysis," *Cognitive Neurodynamics*, vol. 15, no. 3, pp. 369–388, 2021.

[22] M. M. Rahman, A. K. Sarkar, M. A. Hossain et al., "Recognition of human emotions using EEG signals: a review," *Computers in Biology and Medicine*, vol. 136, p. 104696, 2021.

[23] Z. Wan, R. Yang, M. Huang, N. Zeng, and X. Liu, "A review on transfer learning in EEG signal analysis," *Neurocomputing*, vol. 421, pp. 1–14, 2021.

[24] X. Geng, D. Li, H. Chen, P. Yu, H. Yan, and M. Yue, "An improved feature extraction algorithms of EEG signals based on motor imagery brain-computer interface," *Alexandria Engineering Journal*, vol. 61, no. 6, pp. 4807–4820, 2022.

[25] M. C. Guerrero, J. S. Parada, and H. E. Espitia, "EEG signal analysis using classification techniques: logistic regression, artificial neural networks, support vector machines, and convolutional neural networks," *Heliyon*, vol. 7, no. 6, p. e07258, 2021.

[26] M. Woodbright, B. Verma, and A. Haidar, "Autonomous deep feature extraction based method for epileptic EEG brain seizure classification," *Neurocomputing*, vol. 444, pp. 30–37, 2021.

[27] M. S. Safi and S. M. M. Safi, "Early detection of Alzheimer's disease from EEG signals using Hjorth parameters," *Biomedical Signal Processing and Control*, vol. 65, p. 102338, 2021.

[28] S. Murugan, C. Venkatesan, M. G. Sumithra et al., "DEMNET: a deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images," *IEEE Access*, vol. 9, pp. 90319–90329, 2021.

[29] K. Li, J. Wang, S. Li et al., "Feature extraction and identification of Alzheimer's disease based on latent factor of multi-channel EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1557–1567, 2021.

[30] N. Deepa and S. P. Chokkalingam, "Optimization of VGG16 utilizing the Arithmetic Optimization Algorithm for early detection of Alzheimer's disease," *Biomedical Signal Processing and Control*, vol. 74, p. 1034, 2022.

[31] M. Cejnek, O. Vysata, M. Valis, and I. Bukovsky, "Novelty detection-based approach for Alzheimer's disease and mild cognitive impairment diagnosis from EEG," *Medical, & Biological Engineering & Computing*, vol. 59, no. 11, pp. 2287–2296, 2021.

[32] Y. Chen, M. R. Strickland, A. Soranno, and D. M. Holtzman, "Apolipoprotein E: structural insights and links to Alzheimer disease pathogenesis," *Neuron*, vol. 109, no. 2, pp. 205–221, 2021.

[33] A. Kaur, S. Jain, and S. Goel, "Sandpiper optimization algorithm: a novel approach for solving real-life engineering problems," *Applied Intelligence*, vol. 50, no. 2, pp. 582–619, 2020.

[34] J. C. W. Lin, Y. Shao, Y. Djenouri, and U. Yun, "ASRNN: a recurrent neural network with an attention model for sequence labeling," *Knowledge-Based Systems*, vol. 212, p. 6548, 2021.

[35] S. Siuly, Ö. F. Alçin, E. Kabir et al., "A new framework for automatic detection of patients with mild cognitive impairment using resting-state EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 9, pp. 1966–1976, 2020.

Research Article

Feasibility and Safety of Mesocolon Excision with Medical Imaging: A Systematic Review and Meta-Analysis

Chengkui Liu,[1] Qiang Zhao,[1] and Zhichao Song [2]

[1]*Department of Gastrointestinal Surgery, Zibo Central Hospital, Zibo, Shandong Province 255000, China*
[2]*Department of Anorectal Surgery, The First Hospital of Zibo City, No. 4 Emei Shandong Road, Zibo City, Shandong Province 255200, China*

Correspondence should be addressed to Zhichao Song; 15000440229@xs.hnit.edu.cn

The development of new technologies based on electronic intelligent images is a very active research and promotion of new technologies in recent years. This article mainly summarizes the basic concept, development, and technology of electronic intelligent imaging technology, as well as the research, promotion, and application of electronic intelligent imaging technology in clinical treatment. It especially emphasizes the practicality and application of electronic intelligent imaging technology in the current clinical operation process and conducts a meta-analysis of the current mesorectal excision, so as to provide more scientific and professional guidance for clinical surgery. The results of the meta-analysis showed that 3291 documents were initially obtained and duplicate documents were deleted by searching for keywords in mesocolon excision. We excluded 2399 subjects and articles whose interventions did not meet the inclusion criteria of this study after reading the title and abstract. Then, we obtained 892 papers that may meet the inclusion criteria through preliminary screening. We further optimized the search strategy based on selection criteria and data integrity filtering principles and finally determined 111 references. 100 articles that did not meet the requirements were excluded, and 11 articles were finally included for meta-analysis. Medical imaging can effectively improve the therapeutic effect of mesocolon excision and reduce the occurrence of complications. Therefore, it is very important to combine medical intelligent images for preoperative evaluation, and the development of the combination of surgical treatment and medical images should not be underestimated in the future.

1. Introduction

With the development of medical imaging technology, people combine clinical treatment in different fields with medical imaging more closely [1]. At the same time, the combination of medical imaging and surgical diagnosis and treatment can effectively improve the postoperative rehabilitation effect and quality of life of patients [2]. The purpose of this study is to make statistics on the postoperative complications of patients undergoing mesocolon excision and analyze the treatment differences between laparoscopic and open surgery. We try to help medical personnel better perform mesocolon surgery, so as to better develop medical plans.

Total mesocolectomy (TME) was first proposed by a professor head of the United Kingdom in 1988 and gradually developed into a mature surgical treatment scheme. It is believed that rectal cancer surgery should not only remove the focus but also comprehensively remove the diseased tissues including lymph nodes, so as to improve the rehabilitation effect of patients [3]. With the development and wide application of TME, the surgical effect of rectal cancer has been significantly improved. Like rectal cancer, lymph node metastasis is the first metastasis of colon cancer. A large number of clinical studies have shown that fully resecting regional lymph nodes can significantly improve the prognosis and quality of life of patients [4, 5].

The appearance of laparoscopic surgery has further improved the treatment effect of related surgery. Laparoscopic surgery has been applied in Russia since 1901. China successfully completed the first laparoscopic surgery in 1991. In 1901, Ott, a gynecologist in Petersburg, Russia, made a small incision in the front wall of the abdomen, inserted a

speculum into the abdominal cavity, and reflected the light into the abdominal cavity with a head mirror to examine the abdominal cavity. In the same year, German surgeon Kelling inserted a cystoscope into the abdominal cavity of the dog for examination and called this examination laparoscopic endoscopy. In 1938, the Hungarian surgeon Veress introduced an air injection needle, which can safely make pneumothorax; During pneumoperitoneum, the needle tip can be prevented from damaging the viscera under the needle [6]. The idea of making pneumoperitoneum with a compromise safety puncture needle is generally accepted and has been used up to now. Kalk, a German gastroenterologist, invented a lens system with a direct forward squint of 135° and first advocated the use of double trocar puncture needle technology in 1929 [7]. Relevant research found that nearly 1/3 of gynecological operations in the Cedars snai Medical Center in Los Angeles were diagnosed or treated with laparoscopy.

This article reviews the concept, development, research, and application of mesocolon excision and laparoscopy. Meta-analysis was used to analyze the complications of patients after mesocolon excision in China, especially the practical application effect of laparoscopy in surgical treatment. The auxiliary treatment of medical imaging technology can reduce the inflammation of patients to a certain extent. Advanced laparoscopic technology can reduce the external trauma during the operation and improve the operation quality to a certain extent.

In order to analyze and summarize the application of mesocolon excision in clinical practice in China, we searched all relevant nursing disciplines in China National Knowledge Network (CNKI) and web of science. Meta-analysis articles were retrieved from the establishment of the database to August 2022, and the retrieved documents were analyzed. Finally, a total of 892 articles were included in the analysis. By deleting duplicates, a total of 111 articles were included in the analysis, accounting for 12.44% of the total included articles. Through the final screening, a total of 11 articles were obtained. Improve the quality of literature screening by increasing the year of analysis. However, the number of high-quality meta-analysis articles is relatively small. On the one hand, meta-analysis can be widely used in the evaluation of clinical surgical results. On the other hand, Chinese clinical workers should make better use of the best evidence and write high-quality articles.

2. Decision Tree Algorithm and Laparoscopic Image Technology

2.1. Decision Tree Algorithm. The current sample set is divided into two subsample sets by decision tree algorithm using binary recursive segmentation technology. Generally, according to specific segmentation rules, select specific attribute values as branch points, divide the data set into two subsets and repeat this process [8]. When the current subset satisfies the specified segmentation end rule, the growth of the tree stops. After the decision tree is established, the path from the root to the leaf node corresponds to some corresponding logic and rules.

2.1.1. Growth of Decision Tree. In this article, the sample set divided by the CART algorithm refers to the microarray data. The root node is the sample to be divided in the entire microarray dataset, where each gene is the attribute to be selected. And the bifurcation node of the tree is a selected attribute value, and the leaf nodes of the tree represent the class to which the tester belongs [9]. The growth process of the decision tree is the construction process of the classifier.

(1) Error Index. The error index is used to quantify the performance of tree node d branch samples from different classification problems. The error index of this classification tree is often called a mixed subfunction [10].

In the classification tree, the most famous mixture degree function is the Gini index. The Gini index is the most popular and most commonly used division rule, and it can be used well for data containing noise [11]. In the CART algorithm, Gini's indicator is often the default division indicator, because its performance is often the best. It can be used as an error indicator to generate a tree, and it is defined as follows:

$$\phi(H) = -\sum_{L \neq 1} hKhn = 1 - \sum_{L=1}^{L} hL^2. \tag{1}$$

Among them $\sum_{L=1}^{L} hL = 1$, the collection of microarray data samples to be classified can be regarded as the nodes of a classification tree, and the mixing degree function $\varphi(H)$ is used to express the mixing degree index of the tree node d, which is

$$M(d) = \phi(h1, h2, hL). \tag{2}$$

When $L = 2$, $M(d) \in [0, 1/2]$, the smaller its value, the more the two types of samples in the tree node D sample set tend to be one type of sample, that is, the purer.

(2) The Rule of Division. If the mixed order function used to calculate the cost of the node is selected, each time the tree grows, the best branch value that divides the sample in the node into the smallest cost will be found [12]. In the binary tree constructed by the CART algorithm, the amount of change in the degree of mixing caused by the branch is as follows:

$$\Delta M(d) = M(d) - hlM(dl) - hrM(dr), \tag{3}$$

d is the node that is forking; $M(d)$ is the degree of mixing of node d; $M(dl)$ and $M(dr)$ are the mixing degree of left and right branch nodes respectively; hl and hr are the percentages of left and right bifurcated samples in node d, respectively. For the bifurcation of each inner node d, taking the one with the largest amount of change in the degree of mixing among all possible bifurcation modes of d [13].

(3) End of Division. If one of the following conditions is met, the sample segmentation will be stopped:

First, when the samples contained in a certain node belong to the same category, it is already the purest, and the segmentation is stopped

Second, when there are no remaining attributes as the basis for sample division, stop the division

Third, the user sets the maximum depth of the tree in advance to directly limit the growth of the tree and stop the segmentation

2.1.2. Pruning of Decision Tree.

The pruning process of the decision tree aims to find the optimal tree partition and reduce the model overfitting by improving the generalization performance of the decision tree. This is an effective means to ensure that the generated classification tree can match the new data through appropriate training. Therefore, there needs to be a way to stop the growth of the tree at an appropriate time, which is the so-called prepruning method [14]. Among them, one is to set the maximum depth of the tree to directly limit the growth of the tree, and the other is to set the minimum number of records contained in each node. The model stops splitting when the number of records in the node is less than this value [15]. The opposite of prepruning is the postpruning method that allows the tree to grow fully and then prunes it later. For the fully grown tree, selectively pruning upwards to obtain a pruning sequence of a candidate subtree. Then, using an independent test data set or cross-validation method to identify the subtree with the lowest error classification rate from these candidate subtrees as the best subtree [16].

(1) Definition of Pruning. Generally, a tree can be represented by S, and the subtree whose root node is s is represented by S_s, then the pruned subtree S_{s3} is shrunk into an end point $s3$. The pruned tree can be represented as $S - S_{s3}$, and $S - S_{s3} \subset S$, using $\overline{S}$ to represent the set of endpoints in the tree S, and the number of corresponding endpoints is $\overline{S}$. The definition of tree S impurity index is

$$M(S) = \sum_{s \in S} M(s). \tag{4}$$

$M(S)$ is the index of mixing degree of tree node s.

When pruning a decision tree, it needs to pass a cost complexity test first. The cost complexity test is defined as follows:

$$M_\delta(S) = M(S) + \delta|\overline{S}|. \tag{5}$$

In the formula, $M_\delta(S)$ is the linear combination of the cost $M(S)$ of the tree and its complexity. For each given value of δ, the corresponding cost complexity test can always find a minimum subtree $S(\delta)$:

$$M_\delta(S(\delta)) = \min_{S \subset Smax} M_\delta(S). \tag{6}$$

(2) Principle of Pruning. The main idea of gradual pruning: assuming that the tree S has N end points, that is, constantly looking for smaller and smaller tree sequences [17]. θ is the penalty cost of each node: a number that starts from 0 and increases. When $\theta = 0$, there is no penalty for the nodes of the tree, and the cost complexity is measured by the tree that has not started pruning. When θ is increased to a large value, the penalty cost for misclassification can be almost ignored in

the cost complexity measure, and the minimum tree obtained is the smallest and simplest tree, that is, there is only one node [18].

The method of finding the next minimum tree of the tree S is: for each internal node n of the tree S, the value of θ of the next tree S-Ss needs to be obtained, denoted as θ s. This value represents the ratio of the change in the error index before and after the pruning of the current tree to the change in the number of endpoints:

$$\theta_s = \frac{M(s) - M(S_s)}{-1}. \tag{7}$$

In the pruning process of the decision tree, multiple candidate tree sequences will be obtained. To choose the optimal scale tree, it can be measured by two methods [19]. The simplest is to use new sample data to split the candidate tree so that it is easy to pick the tree with the smallest error test.

2.2. Laparoscope.

With the development of modern medical technology, laparoscopy, as an extension of doctors' eyes and hands, has reached the state of "omnipotence" [20]. Laparoscopic treatment can effectively detect abdominal inflammation, ulcers, and benign and malignant tumors. In addition, through laparoscopic adjuvant treatment, the location and scope of organ lesions can be located, and biopsies and color painted specimens can be performed. It can not only effectively diagnose diseases, but also carry out targeted treatment [21].

In the process of obtaining the laparoscope, the 3D scene of the objective world needs to be projected onto the 2D image plane of the camera [22], and this projection can be explained by image conversion. Figure 1 shows a coordinate system related to 3D space scene imaging.

The image coordinate system generally performs image processing in the pixel coordinate system of the image arrangement to indicate the position of the image pixels in the image arrangement, as shown in Figure 2. The projection coordinates of the image plane may perform an image processing process by converting the image coordinate system in the imaging process [23].

If the same-order coordinates of a specific point m in the world coordinate system and the camera coordinate system are $\begin{bmatrix} X & Y & Z & U \end{bmatrix}^N$ and $\begin{bmatrix} x & y & z & u \end{bmatrix}^N$, respectively, the following relationship exists:

$$\begin{bmatrix} x \\ y \\ z \\ u \end{bmatrix} = \begin{bmatrix} M & n \\ O & u \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ U \end{bmatrix}. \tag{8}$$

M is an orthogonal arrangement of $3*3$ units, and n is a $3*1$ translation vector of the image plane coordinate system.

The pixel coordinates of the computer image coordinate system are pixel units. Assuming that the origin of the image plane coordinate system is in the computer image coordinate system, the origin coordinates can be expressed as (a_0, b_0). In the computer image coordinate system, it is usually assumed that the physical size of each pixel of the x'-axis and
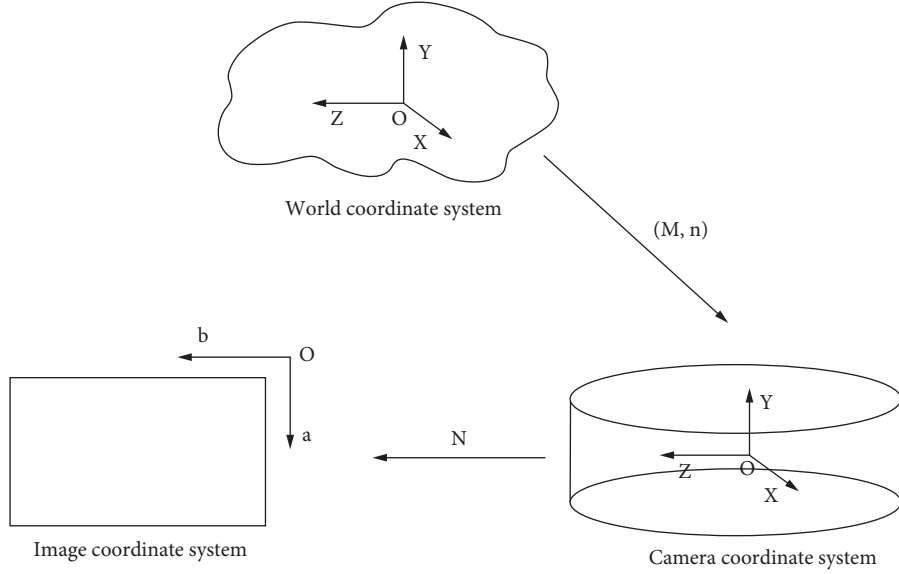
FIGURE 1: Schematic diagram of perspective projection inverted imaging geometry.
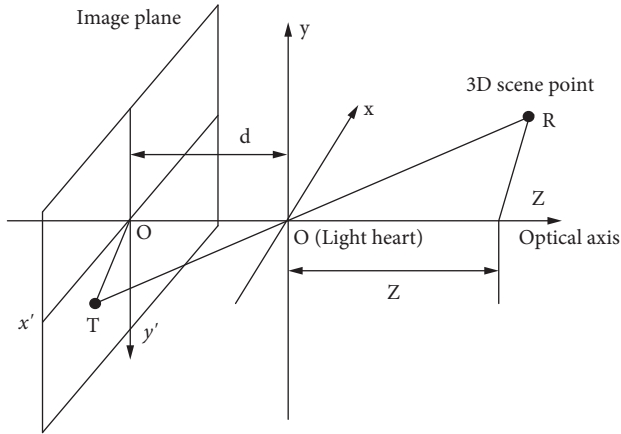
FIGURE 2: Coordinate system.

the y'-axis is fx', fy'. The relationship between the co-ordinates (a, b) of any pixel in the image of the computer image coordinate system and the corresponding coordinates $F(a, b)$ in the image plane is as follows:

$$a = \frac{x'}{fx'} + a_0, \tag{9}$$

$$b = \frac{y'}{fy'} + b_0. \tag{10}$$

Using secondary coordinates, equations (9) and (10) can be expressed as follows:

$$\begin{bmatrix} a \\ b \\ u \end{bmatrix} = \begin{bmatrix} \dfrac{1}{fx'} & 0 & a_o \\ 0 & \dfrac{1}{fy'} & b_0 \\ 0 & 0 & u \end{bmatrix} \begin{bmatrix} x' \\ y' \\ u \end{bmatrix}. \tag{11}$$

The inverse equation of (11) can be expressed as follows:

$$\begin{bmatrix} x' \\ y' \\ u \end{bmatrix} = \begin{bmatrix} fx' & 0 & -a_o fx' \\ 0 & fy' & -b_0 fy' \\ 0 & 0 & u \end{bmatrix} \begin{bmatrix} a \\ b \\ u \end{bmatrix}. \tag{12}$$

2.2.1. Imaging Model. The conversion between a three-dimensional scene and a two-dimensional scene is the most commonly used technology in imaging systems. In the process of converting three-dimensional images into two-dimensional images, most of them will choose to use perspective projection transformation and orthogonal projection transformation [24]. People tend to use perspective projection transformation rather than orthogonal projection transformation in real scenes. Figure 1 shows a schematic diagram of perspective projection inverted imaging geometry.

As the most commonly used imaging model, perspective projection usually reverses the image when performing perspective projection. In order not to produce an inverted image during perspective projection, the image plane will be placed in front of the projection center most of the time. Figure 3 shows a schematic diagram of perspective projection geometry.

In Figure 4, the line of sight RO1 from the point $R(a, b, c)$ of the 3D scene, the vertical line RO3 from the point $R(a, b, c)$ to the z-axis, and O1R1 form a triangle. The line of sight TO3 of point $T(a', b')$ on the image plane, the vertical line from point $T(a', b')$ D to the z-axis, and the z-axis form a triangle. These two triangles are similar triangles, from which the perspective projection equation is obtained:

$$\frac{a'}{a} = \frac{b'}{b} = \frac{d}{c}. \tag{13}$$
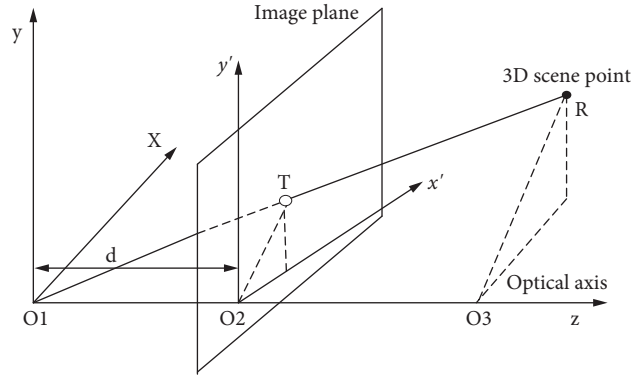
FIGURE 3: Optical flow and splicing.
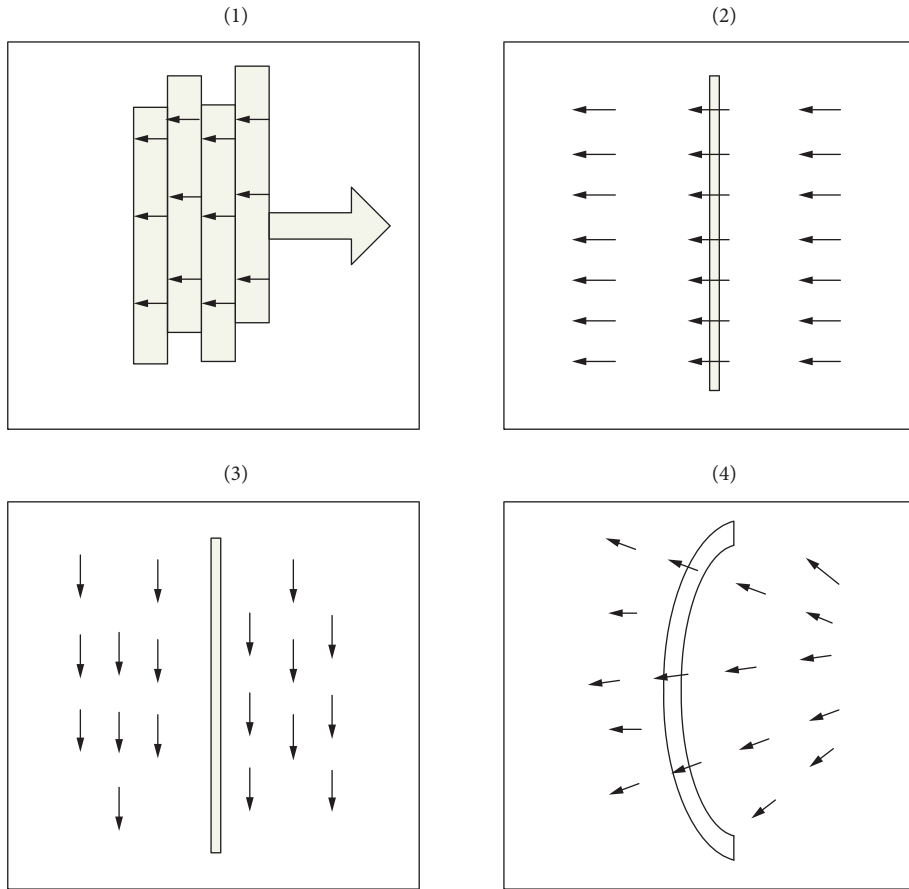
FIGURE 4: Schematic diagram of perspective projection geometry.

The position of the 3D scene point in the image plane is given by the following equation:

$$a' = \frac{d}{c}a,$$
$$b' = \frac{d}{c}b. \qquad (14)$$

In the formula, (a, b, c) represents the coordinates of the scene point R in the camera coordinate system, and (a', b') represents the coordinates of the T point in the plane image.

The above-mentioned projection relationship can be expressed by a homogeneous coordinate equation, and the specific expression is as follows:

$$c\begin{bmatrix} a' \\ b' \\ u \end{bmatrix} = \begin{bmatrix} d & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & u & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ u \end{bmatrix}. \qquad (15)$$

If equations (8) and (12) are substituted into (15), the relationship between the coordinates of the 3D scene point R

represented by the world coordinate system and the coordinates of the projection point T in the computer image coordinate system can be obtained:

$$z \begin{bmatrix} a \\ b \\ u \end{bmatrix} = \begin{bmatrix} \dfrac{1}{fx'} & 0 & 0 \\ 0 & \dfrac{1}{fy'} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} d\,0\,0\,0 \\ 0\,d\,0\,0 \\ 0\,0\,u\,0 \end{bmatrix} \begin{bmatrix} M\,n \\ O\,u \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ U \end{bmatrix} = E \begin{bmatrix} X \\ Y \\ Z \\ U \end{bmatrix}. \quad (16)$$

Among them, E is called the projection matrix.

2.2.2. Mosaic of Endoscopic Images. Generally speaking, the splicing of endoscopic images mainly includes several steps of endoscopic image acquisition, image preprocessing, image registration, reprojection model selection, and image fusion [25].

(1) Manifold Stitching. Figure 3 shows the general process of endoscopic image stitching.

If the displacement of two adjacent images is small, the two images can be regarded as the relationship of affine motion, so there is the following calculation formula:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x_i - x_{i-1} \\ y_i - y_{i-1} \end{bmatrix} = \begin{bmatrix} m + nx_i + vy_i \\ z + ax_i + by_i \end{bmatrix}. \quad (17)$$

Then, the transformation relationship between image T_{i-1} and image T_i is

$$T_{i-1} = M_i T_i, \quad (18)$$

where M_i is

$$M_i = \begin{bmatrix} 1-n & -v & -m \\ -a & 1-b & -z \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \cos\theta * sx & -\sin\theta * shx & fx \\ \sin\theta * shy & \cos\theta * sy & fy \\ 0 & 0 & u \end{bmatrix}. \quad (19)$$

According to Figure 5, to achieve cutting, it needs to find the line $X\,(x, y) = 0$ perpendicular to all optical flows. Since the normal vector of $X\,(x, y)$ is perpendicular to X, we can get the following equation:

$$\begin{pmatrix} \dfrac{\partial M}{\partial x} \\ \dfrac{\partial M}{\partial y} \end{pmatrix} = L \begin{pmatrix} a \\ b \end{pmatrix} = L \begin{pmatrix} m + nx + vy \\ z + ax + by \end{pmatrix}. \quad (20)$$

When $v = a$, the solution of M can be obtained from equation (20) as follows:

$$0 = M\,(x, y) = mx + zy + \frac{n * x^2}{2} + (v + a)x * y + \frac{b * y^2}{2} + Y. \quad (21)$$
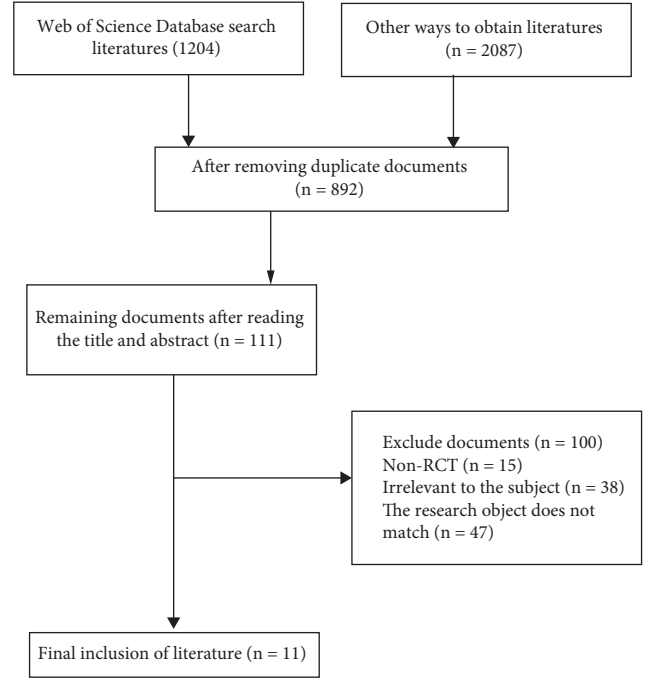
FIGURE 5: Document retrieval process.

3. Method

3.1. Search Strategy. The computer searches the web of science and CNKI. The search period is from the establishment of the database to August 2022. The search adopts the combination of free words and subject words and uses the literature tracking method to find relevant literature. The search terms in the English database include "mesocolon excision" and "laparoscopy"; The search terms in the Chinese database include "mesocolon excision," "laparotomy" and "laparoscopy."

3.2. Literature Inclusion and Exclusion Criteria. The inclusion criteria are as follows: ① the study design type is a randomized controlled trial or quasirandomized controlled trial. ② The study subjects in the literature were patients with mesocolon excision. ③ The surgical methods used by the experimental group in the literature include laparoscopic assisted treatment and Da Vinci intelligent robot assisted ④ the outcome indicators include the number and situation of postoperative complications.

Exclusion criteria are as follows: ① documents without full text, incomplete information or unable to extract data. ② Duplicate publication. ③ Minutes of meeting.

3.3. Literature Screening and Data Extraction. Two reviewers independently screened the preliminary literature strictly according to the inclusion and exclusion criteria. After excluding the trials that obviously did not meet the inclusion criteria, the full text of the potentially relevant literature was analyzed. The latter two reviewers cross-checked the results of the included studies and decided on the divergent studies

TABLE 1: Main features of the study.

Author (Year)/Country	Design	Type of study	Number of patients in the laparoscopic group	Number of patients in the open surgery group	Bias risk grade
Bian (2020)/China	RCT	Randomized trial	40	40	B
Chao (2022)/China	RCT	Randomized trial	60	60	B
Cui (2017)/China	RCT	Randomized trial	85	85	B
Fan (2017)/China	RCT	Randomized trial	40	40	B
Lin (2022)/China	RCT	Randomized trial	30	30	Be
Quan (2017)/China	RCT	Randomized trial	47	43	B
Shen (2020)/China	RCT	Randomized trial	45	78	B
Sun (2018)/China	RCT	Randomized trial	50	50	B
Wang (2019)/China	RCT	Randomized trial	51	51	B
Yu (2018)/China	RCT	Randomized trial	102	116	B
Zhang (2017)/China	2-Arm RCT	Randomized trial	26	26	B
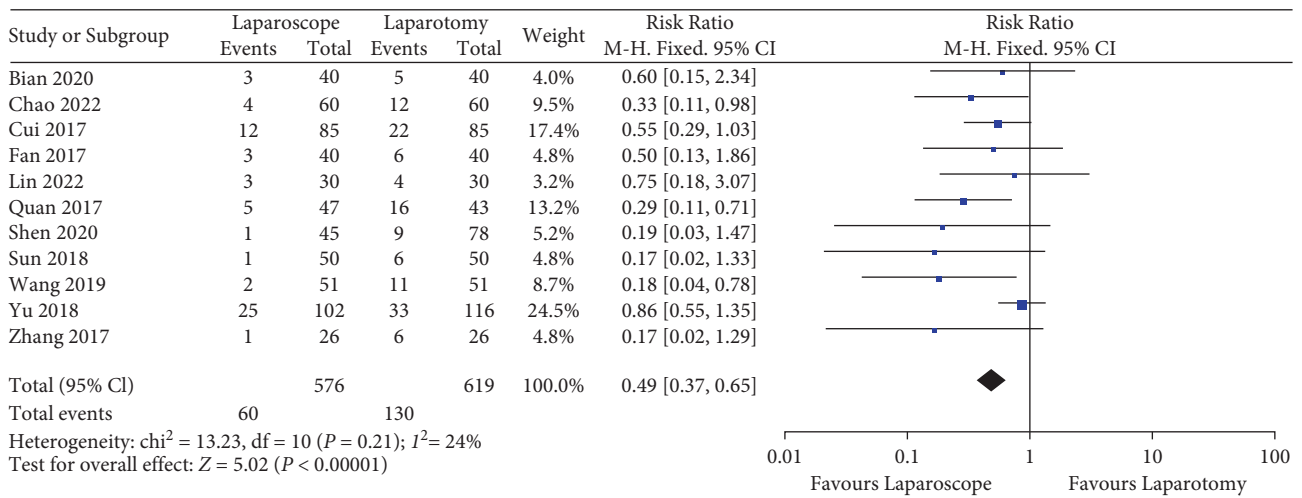
FIGURE 6: Meta-analysis forest plots expressed in ≤60 years old and >60 years old.

through discussion or with the third investigator. For those with incomplete reports of included literature information, they can be supplemented by contacting the authors.

3.4. Meta-Analysis. This study used RevMan5.3 software for meta-analysis. Meta-studies generally have heterogeneity, which can be divided into methodological heterogeneity, clinical heterogeneity, and statistical heterogeneity. The heterogeneity of methods is caused by differences in research design and quality. Clinical heterogeneity is caused by differences in participants, interventions, and endpoint indicators. Statistical heterogeneity is based on data and overlaps with the confidence interval between studies. Correlation, the degree of coincidence is large, and the heterogeneity is small, on the contrary, the heterogeneity is large. If meta-analysis includes studies with large heterogeneity, it will reduce the credibility of the conclusions. At the same time, the existence of heterogeneity among various studies determines the choice of the combined effect size model.

4. Results

4.1. Literature Search Results. The preliminary inspection obtained 3291 documents. After reading the title and abstract, 2399 articles whose subjects and interventions did not meet the inclusion criteria of this study were excluded, and 111 papers that might meet the inclusion criteria were obtained through preliminary screening. After reading the full text and tracing the references of the included documents, no other documents related to this study were found. Further reading was screened according to inclusion and exclusion criteria and data integrity. 100 articles were excluded and 11 articles were included for meta-analysis. The document retrieval process is shown in Figure 5.

4.2. Quality of Included Studies. The methodological quality evaluations of the included 11 studies were all B-level.

4.3. Meta Analysis Results. This study analyzed the application effect of laparoscopy in mesocolon excision and the
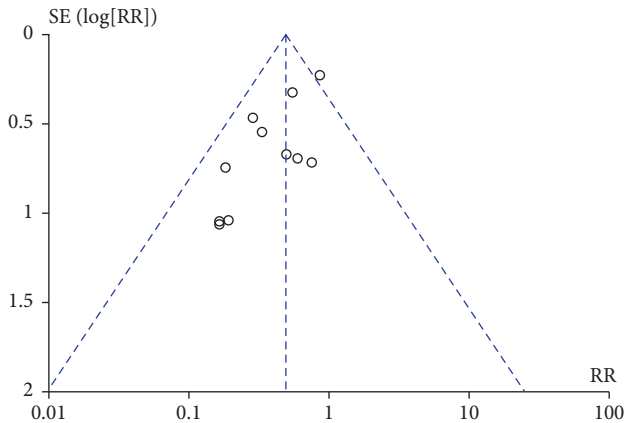
FIGURE 7: Funnel plot of comparison: 1 laparoscope vs. laparotomy.

postoperative complications of patients with two surgical methods.

4.4. Theoretical Knowledge. Eleven studies evaluated the application effect of virtual reality technology at the theoretical knowledge level. The results showed that the heterogeneity of the included studies was low ($P = 0.21$, $I2 = 24\%$), so the fixed effect model was adopted. The comprehensive results showed that compared with open surgery, the use of laparoscopy can effectively reduce the probability of postoperative complications [$Z = 5.02$, 95% CI (0.37, 0.65), $P < 0.01$].

4.5. Satisfaction Survey. The evaluation process of surgical effect is based on certain clinical indicators or the quality of life of patients after surgery. The occurrence of postoperative complications can effectively judge the quality of surgery. We used meta-analysis to make statistics on the literature used in this article. The results are shown in Table 1.

4.5.1. Analysis of Experimental Results. A meta-analysis of 11 articles was performed, showing laparoscopic and open surgery in 90 patients, respectively, as shown in Figure 6.

Figure 7 shows the funnel chart results of 11 meta-analyses. The funnel chart results show that the distribution of the literature in this study is balanced and the reliability of the results is high.

5. Discussion and Conclusion

5.1. Advantages of Laparoscopic Surgery. Through meta-analysis, this study concluded that laparoscopy and electronic imaging technology can be fully combined with surgical clinical treatment, which can effectively improve the quality of life of patients. Modern clinical treatment should pay attention to the development and application of medical imaging technology, especially the diagnosis and treatment of complex surgery. Therefore, the clinical medical college should increase the investment in emerging technologies. The organic combination of electronic imaging technology

and clinical practice can play an important role in the future clinical surgery practice.

5.2. Influence of Laparoscopic Treatment on Clinical Operation Effect. Generally speaking, electronic intelligent imaging technology is currently a relatively active technology, and clinical practice makes the future development of this technology full of vitality. Although significant results have been achieved in the practical application, there is still a more general gap between the requirements of clinical application. The development of this technology in China is relatively late, but with the gradual attention of relevant departments to virtual reality technology, the development of the Da Vinci surgical robot will be more prominent in clinical medicine, which will bring new development opportunities for future clinical treatment. At the same time, this meta-analysis has the following limitations: the number of studies included is small and the sample size is small. The results of the meta-analysis still need to be further verified by more large sample multicenter studies.

Data Availability

The data used to support the study are included in the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Chengkui Liu, Qiang Zhao these authors contributed equally to this work.

References

[1] M. Arif and G. Wang, "Fast curvelet transform through genetic algorithm for multimodal medical image fusion," *Soft Computing*, vol. 24, no. 3, pp. 1815–1836, 2020.

[2] S. Sajedi, H. Sabet, and H. S. Choi, "Intraoperative biophotonic imaging systems for image-guided interventions," *Nanophotonics*, vol. 8, no. 1, pp. 99–116, 2018.

[3] N. P. West, H. Kobayashi, K. Takahashi et al., "Understanding optimal colonic cancer surgery: comparison of Japanese D3 resection and European complete mesocolic excision with central vascular ligation," *Journal of Clinical Oncology*, vol. 30, no. 15, pp. 1763–1769, 2012.

[4] W. W. Dzwierzynski, "Complete lymph node dissection for regional nodal metastasis," *Clinics in Plastic Surgery*, vol. 37, no. 1, 2009.

[5] T. Konishi and Y. N. You, "Complete mesocolic excision and extent of lymphadenectomy for the treatment of colon cancer," *Surgical Oncology Clinics of North America*, vol. 31, no. 2, pp. 293–306, 2022.

[6] K. Jackson and R. Rauck, "Proper needle size for performing cervical epidural injections," *Anesthesiology*, vol. 75, no. 1, p. 166, 1991.

[7] G. Schaller, M. Kuenkel, and B. C. Manegold, "The optical "Veress-needle"-initial puncture with a minioptic," *Endoscopic Surgery & Allied Technologies*, vol. 3, no. 1, pp. 55–57, 1995.

[8] J. Chan, B. Chan, H. L. Ho, K. M. Chan, P. G. Kan, and HS. Lam, "The neonatal resuscitation algorithm organized cart is more efficient than the airway–breathing–circulation organized drawer: a crossover randomized control trial," *European Journal of Emergency Medicine*, vol. 23, no. 4, pp. 258–262, 2016.

[9] S. Celik, B. Sogut, T. Sengul, E. Eyduran, and AY. Sengul, "Usability of CART algorithm for determining egg quality characteristics influencing fertility in the eggs of Japanese quail," *Revista Brasileira de Zootecnia*, vol. 45, no. 11, pp. 645–649, 2016.

[10] A. Kocakusak, A. Kocakusak, and A. Kocakusak, "Does Ramadan fasting contribute to the increase of peptic ulcer perforations?" *European Review for Medical and Pharmacological Sciences*, vol. 21, no. 1, pp. 150–154, 2017.

[11] Y. J. Yang, C. S. Bang, S. P. Shin et al., "Clinical characteristics of peptic ulcer perforation in Korea," *World Journal of Gastroenterology*, vol. 23, no. 14, pp. 2566–2574, 2017.

[12] D. Musa, "Peptic ulcer disease and non-steroidal anti-inflammatory drugs," *Australian Prescriber*, vol. 40, no. 3, pp. 91–93, 2017.

[13] G. Sgourakis, G. Chatzidakis, A. Poulou et al., "High-dose vs. Low-dose Proton Pump Inhibitors post-endoscopic hemostasis in patients with bleeding peptic ulcer. A meta-analysis and meta-regression analysis," *Turkish Journal of Gastroenterology*, vol. 29, no. 1, pp. 22–31, 2018.

[14] H. Kato, R. Pattison, and S. Bhandari, "Inverse relationship between *Helicobacter pylori* infection and asthma in US adults with peptic ulcer disease," *Indian Journal of Gastroenterology*, vol. 36, no. 2, pp. 158-159, 2017.

[15] T. Shou-Jiang, C. Dimitrios, Z. Simon et al., "Wireless capsule endoscopy for obscure gastrointestinal bleeding: a single-centre, one-year experience," *Canadian Journal of Gastroenterology*, vol. 18, no. 9, pp. 559–565, 2016.

[16] F. Cardona, P. Salanski, and M. Chmielewski, "Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline," *Endoscopy*, vol. 50, no. 4, pp. 423–446, 2018.

[17] R. D. Acosta, N. S. Abraham, V. Chandrasekhara et al., "The management of antithrombotic agents for patients undergoing GI endoscopy," *Gastrointestinal Endoscopy*, vol. 83, no. 1, pp. 3–16, 2016.

[18] S. Goulas, K. Triantafyllidou, S. Karagiannis et al., "Capsule endoscopy in the investigation of patients with portal hypertension and anemia," *Canadian Journal of Gastroenterology*, vol. 22, no. 5, pp. 469–474, 2008.

[19] A. E. Mohamad, D. G Anton, C. Vinay et al., "The role of endoscopy in the evaluation and management of patients with solid pancreatic neoplasia," *Gastrointestinal Endoscopy*, vol. 83, no. 1, pp. 17–28, 2016.

[20] C. S. Sydney, "Preventing ulcer rebleeding: the role of second-look endoscopy," *Canadian Journal of Gastroenterology*, vol. 13, no. 5, pp. 409–411, 2016.

[21] M. G. Kundu and J. Harezlak, "Regression trees for longitudinal data with baseline covariates," *Biostatistics & epidemiology*, vol. 3, no. 1, pp. 1–22, 2018.

[22] S. H. Mirhashemi, P. Haghighat jou, F. Mirzaei, and M. Panahi, "The study of environmental and human factors affecting aquifer depth changes using tree algorithm," *International journal of Environmental Science and Technology*, vol. 17, no. 3, pp. 1825–1834, 2020.

[23] A. S. Badi Albaqawi, N. M. Abo El-Fetoh, R. F. Abdullah Alanazi et al., "Profile of peptic ulcer disease and its risk factors in Arar Northern Saudi Arabia," *Electronic Physician*, vol. 9, no. 11, pp. 5740–5745, 2017.

[24] A. Araújo, S. Barbeiro, E. Cuesta, and A. Duran, "Cross-diffusion systems for image processing: I. The linear case," *Journal of Mathematical Imaging and Vision*, vol. 58, no. 3, pp. 447–467, 2017.

[25] V. Brea, D. Ginhac, F. Berry, and R. Kleihorst, "Special issue on advances on smart camera architectures for real-time image processing," *Journal of Real-Time Image Processing*, vol. 14, no. 3, pp. 635-636, 2018.

Research Article

Design and Development of a Big Data Platform for Disease Burden Based on the Spark Engine

Chengcheng Li,[1] **Jing Gao,**[1] **Qingwei Pan,**[2] **Zhihua Zhou,**[1] **Yue Yang,**[3] **and Shangcheng Zhou** [1]

[1]*School of Public Health and Management, Guangzhou University of Chinese Medicine, Guangzhou 510006, China*
[2]*College of Physical Education and Health, Guangxi Medical University, Nanning 530021, China*
[3]*School of Life Science and Technology, Inner Mongolia University of Science and Technology, Baotou 014010, China*

Correspondence should be addressed to Shangcheng Zhou; zhoushangcheng@gzucm.edu.cn

Objective. This study attempts to build a big data platform for disease burden that can realize the deep coupling of artificial intelligence and public health. This is a highly open and shared intelligent platform, including big data collection, analysis, and result visualization. *Methods.* Based on data mining theory and technology, the current situation of multisource data on disease burden was analyzed. Putting forward the disease burden big data management model, functional modules, and technical framework, Kafka technology is used to optimize the transmission efficiency of the underlying data. This will be an efficient and highly scalable data analysis platform through embedding embedded Sparkmlib in the Hadoop ecosystem. *Results.* With the concept of "Internet + medical integration," the overall architecture design of the big data platform for disease burden management was proposed based on the Spark engine and Python language. The main system composition and application scenarios are given at four levels: multisource data collection, data processing, data analysis, and the application layer, according to application scenarios and use requirements. *Conclusion.* The big data platform of disease burden management helps to promote the multisource convergence of disease burden data and provides a new path for the standardized paradigm of disease burden measurement. Provide methods and ideas for the deep integration of medical big data and the formation of a broader standard paradigm.

1. Background

The explosive prosperity of Internet of things technology and artificial intelligence has promoted profound changes in the industry worldwide [1]. Information resources are gradually becoming one of the most essential strategic resources. This large-scale, fast-growing, diverse data structure and multidimensional value density data set is often called big data [2]. How to identify and extract the key features of information in a timely, accurate, and efficient manner from the data ocean and conduct rapid analysis has become one of the hot issues [3, 4].

There are many positive explorations in the construction of big data platforms in many fields around the world. A study from Spain explored the construction of a social media big data platform, which can monitor early signs of depression in 2020 [5]. In the field of transportation, some scholars have realized the dynamic benefits of monitoring of traffic violations through a big data platform [6]. Moreover, the education big data platform achieved the informational management of student courses and effectively improved the teaching quality in the field of education [7]. Some scholars have built a big data platform based on the lambda architecture and applied it in the energy field to realize the visual monitoring of residential power consumption and effectively improve energy utilization efficiency [8]. Not only that, but the big data platform based on the lambda architecture can also dynamically monitor and analyze marine activities and integrate various parameters [9]. In addition to it, cloud computing technology provides new solutions and

computing potential for the storage and management of large-scale data, relying on the scientific Earth big data platform [10].

Big data has wider application scenarios and development potential in the field of medical research. The United States promotes the opening and sharing of big data to the medical industry to realize technological innovation. This is a strategic plan to encourage multinational pharmaceutical enterprises to deeply explore the hidden application value in the massive medical information by implementing the "big data research and development plan" [11]. On this basis, multinational pharmaceutical enterprises try to achieve accurate drug development by deeply integrating big data technology with precision medicine [12]. These highly personalized precision treatment schemes have provided great help to further complete the "human genome project" [13]. In 2015, the United States proposed to establish a global open cancer genomics database and named it the "moon landing project" for cancer [14]. The United States has realized the construction of local full-coverage medical big data through 12 electronic case data centers, 9 medical data centers, and 8 biological information databases in 2018 [15]. Meanwhile, Korean doctors have obtained a large amount of data to implement precision medicine by creating a big data platform for cancer [16]. The application of big data medical platforms can reduce neonatal mortality and disability rates effectively [17]. This is a rapidly developing trend in medical big data research around the world.

China has introduced a series of favorable policies to promote technological development and the industrial landing of big data. In 2016, China issued guiding opinions on promoting and standardizing the application and development of big data in health care, aiming to promote the "healthy China 2030" plan to realize the medical mode of "Internet +" for the whole industry chain [18]. Meanwhile, policymakers respond to national strategic needs by innovating the talent training mode and strengthening the talent training of cross-disciplinary fields such as artificial intelligence and big data. In the "13th five-year plan" for the development of national population health informatization in 2017, it is clearly pointed out that the health care big data service system in 2021, which is composed of regional medical big data centers and supporting clinical data analysis centers, will be initially established and issued by the China Health Commission [19]. The "14th five-year plan" for national informatization in 2021 further emphasizes the exploration of the application effect of artificial intelligence in intelligent clinical assistant diagnosis and treatment, intelligent public health service, and artificial intelligence-assisted drug research and development, and emphasizes the key role of the construction of a scientific research big data platform in the development of big data [20]. Codelot, the national life big data platform, was established in Shenzhen with strong policy support and can provide various functions, including gene detection, biological information analysis, and data mining. Chinese scholars have developed a medical imaging platform that relies on cloud computing technology to enhance the diagnostic efficiency of obstetric imaging [21]. The medical big data platform has also played an active role in the rapid gene diagnosis of the digestive department in clinical practice [22].

Disease burden data has the characteristics of high capacity, rapid growth, diverse types, and low value density [23]. Some scholars have examined a lot of high-quality research in the field of disease burden [24–26]. Therefore, the technological breakthrough of big data in disease burden research is very important. At present, the research on disease burden focuses more on manual analysis using machine learning algorithms. Of course, there is a lack of research on the intelligent extraction of massive medical data [27]. At present, there are few unified big data platforms for disease burden research. Hadoop architecture provides us with inspiration and ideas for our research based on Java and Python languages [28].

This study attempts to solve the following problems from the perspective of artificial intelligence: (1) the big data platform for disease burden is designed and created based on the Hadoop distributed computing framework. (2) Optimize the calculation efficiency and accuracy based on the Spark engine in the traditional distributed computing framework. (3) We try to realize the visual design of the analysis results in order to provide an analysis paradigm for the related research on disease burden in the future.

2. Methods

This study attempts to build a big data platform for disease burden analysis based on the Spark engine, which can achieve the whole process of data dynamic capture, storage, analysis, detection, and visual output in disease burden analysis. It can further optimize the efficiency of machine learning algorithms by embedding Spark into Hadoop [29, 30]. There is a wide gap between Hadoop and Spark in the actual construction process, as shown in Table 1.

There are many similarities between Spark and Hadoop, both of which are open-source cluster computing environments. We call it Sparkmlib to realize real-time calculation, access tracking, and anomaly detection of disease burden data.

2.1. Workflow of Spark Streaming. A discretized stream is the data over time, and it is also a sequence composed of the RDD of each time interval [31]. DStreams will be formed from multiple input sources such as Flume, Kafka, or HDFS, as shown in Figure 1.

Specifically, Spark streaming treats streaming computing as a series of continuous small-scale batch processing. Data will be divided into small batches by reading data from different input sources, with the creation of new batches occurring at uniform intervals. A new batch will be formed at the beginning of each time interval. The data received during the interval will be added to the batch. The batch will stop growing at the end of the time interval.

2.2. Big Data Measurement Paradigm of Disease Burden. The design of the mlib algorithm package based on Spark is simple [32]. First, the data is expressed in the form of an RDD, and then various algorithms are called on the distributed dataset. We try to embed JOINTPOINT software

Table 1: Technical differences between Spark and Hadoop.

	Hadoop	Spark
Type	Basic platform, including calculation, storage, and scheduling	Pure distributed computing tools
Scene	Mass data batch processing (disk iterative calculation)	Massive data batch processing (memory iterative calculation, interactive calculation), massive data stream calculation
Price	Low	High
Programming paradigm	MAP + REDUCE	RDD is a DAG directed acyclic graph
	API level is relatively low, and algorithm adaptability is poor	The API is top-level and easy to use
Data storage structure	The calculation result is on the HDFS disk with a large delay	RDD intermediate operation results are stored in memory with a small delay
Operation mode	Tasks are maintained in process mode, and the task starts slowly	Tasks are maintained in a threaded manner, with fast task startup, and can be created in batch to improve the parallel ability
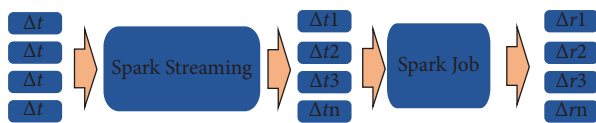
Figure 1: Workflow of Spark streaming.

and DISMOD software into our disease burden big data platform. We have built a whole process big data analysis platform based on Spark engine in this study. The algorithm is shown in Figure 2.

Apache Hadoop Yarn is an ideal Hadoop resource manager [33]. As a general resource management system, it can uniformly schedule applications on the platform. The advantages of strong compatibility can bring great benefits to the cluster. Resource management and scheduling are realized by creating application managers and global managers of MapReduce traditional applications for HDFS [34]. As a supplement to MapReduce, hive improves the ability for rapid development of big data platforms and reduces the difficulty of building big data platforms through SQL-like syntax [35].

In general, we split the big data analysis algorithm into the following four steps, as shown in Figure 3.

2.3. Feature Extraction Algorithm

2.3.1. TF-IDF Algorithm. TF-IDF (term frequency-inverse document frequency) is a classic weighting technique for information retrieval and text mining [36].

TF represents the probability of a keyword's occurrence in the text. Normalization can prevent deviations in text mining. The formula is as follows:

$$tf_{ij} = \frac{n_{i,j}}{\sum_k n_{i,j}}, \tag{1}$$

where $n_{i,j}$ is the number of times the word appears in the file i, j.

Furthermore, this term can have an excellent ability to distinguish categories in the case of fewer documents of T and large IDF. The calculation method is as follows:

$$idf_i = \log \frac{|D|}{\left|D\{j : t_i \in d_j\}\right|}, \tag{2}$$

where $|D|$ is the entire number of files in the corpus. $\{j : t_i \in d_j\}$ denotes the number of files containing the word t_i (i.e., the number of files with n_i, $j \neq 0$). Usually, $1 + |\{j : t_i \in d_j\}|$ is used to avoid the case that the word is not in the corpus and the denominator is 0. Actually, TF-IDF is better at filtering out common words while retaining important words. The calculation method is as follows:

$$TF - IDF = TF * IDF. \tag{3}$$

2.3.2. FP Growth Algorithm. In order to further optimize the keyword screening strategy, we tried to introduce a FP growth algorithm based on the TF-IDF algorithm. [37] This algorithm finds and recommends high-frequency word pairs by looking at the words used on the Internet. The data consistency of medical electronic cases is weak due to the wide range of sources. Therefore, the FP growth algorithm can comprehensively extract and collect data. The algorithm logic is as follows:

First, building an FP Tree based on a certain data structure, as shown in Figure 4.

FP Tree is not used for simple decision trees, so a class should be created to save each node of the tree. The FP Tree will store the occurrence frequency of item sets. Only when the sets are completely different, the tree will fork.

Second, the conditional pattern library is obtained from the FP Tree, and frequent item sets are mined so as to build a larger set on the basis of a single element itemset. This is an effective way to create a conditional FP Tree, which can repeatedly cycle other single-element items for each frequent item.

3. Model Optimization Algorithm

3.1. ALS (Alternating Least Squares) Algorithm. The Spark platform integrates the ALS algorithm. The optimization of matrix decomposition can be realized quickly by constructing different loss functions [38]. The final task of matrix decomposition is to find two matrices, *P* and *Q*, and
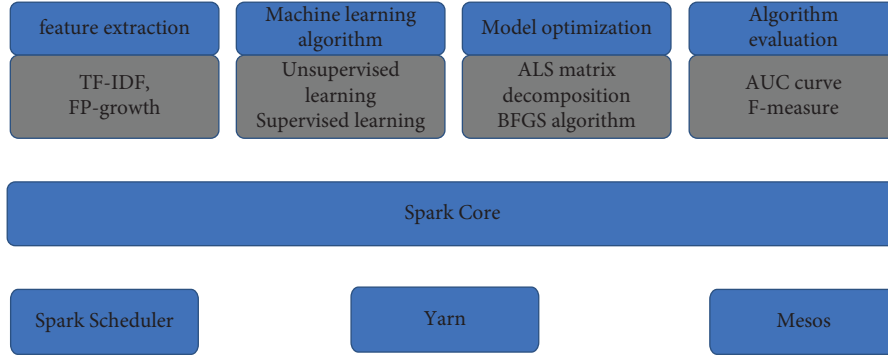
FIGURE 2: Big data measurement paradigm based on Spark engines.
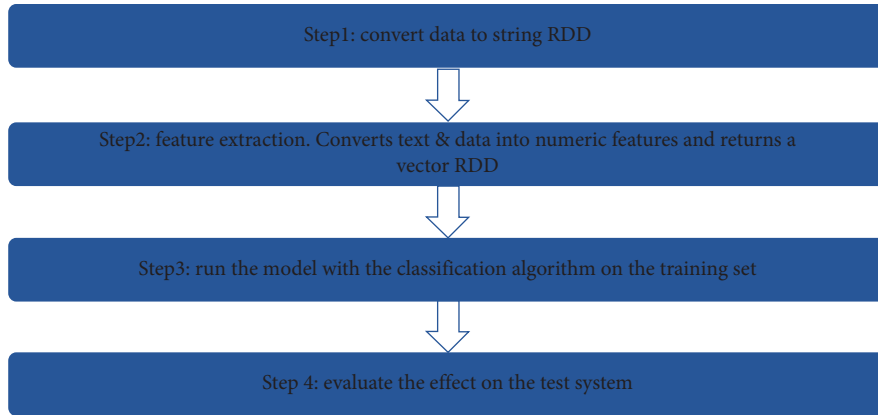
FIGURE 3: Big data analysis steps.

make them approximately equal to the original matrix R after multiplication. The specific algorithm is as follows:

$$R_{m*n} = P_{m*k} \times Q_{n*k}^T, \tag{4}$$

where P and Q are unknown. We assume that Q is known. Therefore,

$$P_{m*k} = R_{m*n} \times Q_{n*k}^{-1}. \tag{5}$$

This means that the R matrix is multiplied by the inverse matrix of the Q matrix, and the result is obtained through

iteration. It is assumed that the solving process is carried out alternately until the error is acceptable.

3.2. L-BFGS Optimization Algorithm. This optimization algorithm evolved from the Newton method and the quasi-Newton method and has been widely used commercially [39]. The specific algorithm is as follows:

Let $f(x)$ be a quadratic differentiable real function, set up again $\chi^{(k)}$ is an estimate of a minimal point of $f(x)$. We expand $f(x)$ into Taylor series at $\chi^{(k)}$ and take the second-order approximation.

$$f(x) \approx \varnothing(x) = f\left(x^{(k)}\right) + \nabla f\left(x^{(k)}\right)^T \left(x - x^k\right) + \frac{1}{2}\left(x - x^{(k)}\right)^T \nabla^2 f\left(x^{(k)}\right)\left(x - x^{(k)}\right), \tag{6}$$

where the middle part of the last item present Hesse matrix of $f(x)$ at $x^{(k)}$. The following formula can be obtained by deriving the abovementioned formula and making it equal to 0:

$$\nabla f\left(x^{(k)}\right) + \nabla^2 f\left(x^{(k)}\right)\left(x - x^k\right) = 0. \tag{7}$$

Assuming that the Hesse matrix is reversible, the iterative formula of the Newton method can be obtained from the abovementioned formula as follows:

$$x^{(k+1)} = x^{(k)} - \nabla^2 f\left(x^{(k)}\right)^{-1} \nabla f\left(x^{(k)}\right), \tag{8}$$

$$x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}, \tag{9}$$

$$d^{(k)} = -\nabla^2 f\left(x^{(k)}\right)^{-1} \nabla f\left(x^{(k)}\right), \tag{10}$$

where λ is the compensation obtained by one-dimensional search, which means
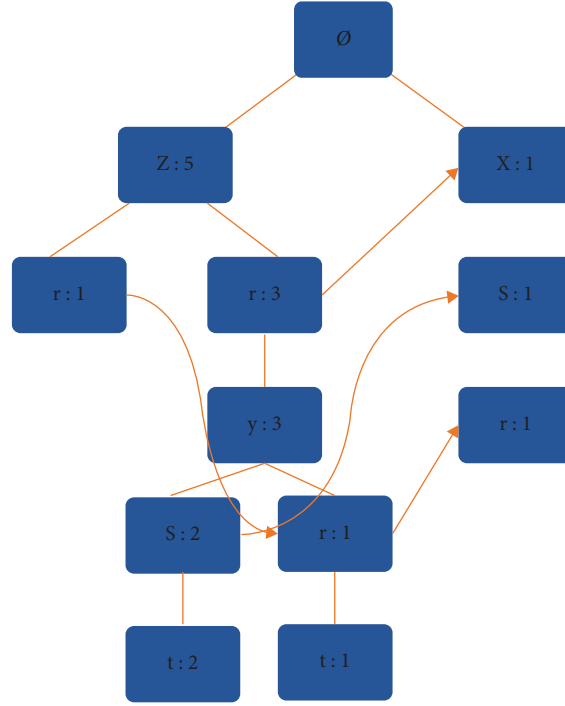
FIGURE 4: FP tree based on certain data structures.

$$f\left(x^{(k)} + \lambda_k d^{(k)}\right) = min f x^{(k)} + \lambda_k d^{(k)}. \tag{11}$$

We try to construct the approximate matrix of the inverse matrix of the Hesse matrix by analyzing the association

between the inverse matrix and the first derivative. Assume that $\chi^{(k+1)}$ is obtained after the k-th iteration. We expand the objective function $f(x)$ into the Taylor series at point $\chi^{(k+1)}$ and take the second-order approximation to obtain

$$f(x) \approx f(x^{(k+1)}) + \nabla f\left(x^{(k+1)}\right)^T (x - x^k) + \frac{1}{2}\left(x - x^{(k+1)}\right)^T \nabla^2 f\left(x^{(k+1)}\right)\left(x - x^{(k+1)}\right). \tag{12}$$

It can be seen that in the vicinity of $\chi^{(k+1)}$,

$$\nabla f(x) \approx \nabla f\left(x^{k+1}\right) + \nabla^2 f\left(x^{k+1}\right)\left(x - x^{(k+1)}\right), \tag{13}$$

$$\nabla f\left(x^{(k)}\right) \approx \nabla f\left(x^{k+1}\right) + \nabla^2 f\left(x^{k+1}\right)\left(x^{(k)} - x^{(k+1)}\right), \tag{14}$$

$$p^{(k)} = x^{(k+1)} - x^{(k)}, \tag{15}$$

$$q^{(k)} = \nabla f\left(x^{(k+1)}\right) - \nabla f\left(x^{(k)}\right), \tag{16}$$

$$q^{(k)} \approx \nabla^2 f\left(x^{(k+1)}\right) p^{(k)}, \tag{17}$$

$$p^{(k)} \approx \nabla^2 f\left(x^{(k+1)}\right) q^{(k)}, \tag{18}$$

$$p^{(k)} = H_k q^{(k)}. \tag{19}$$

Therefore, formula (19) is called the quasi-Newton condition.

When the inverse matrix of the Hesse matrix is a symmetric positive definite matrix, the matrix $H_{(K)}$

satisfying the quasi-Newton condition should also be a symmetric positive definite matrix. We assume that

$$H_{k+1} = H_k + \Delta H_k, \tag{20}$$

$$H_{k+1} = H_k + \frac{\left(p^{(k)} - H_k q^{(k)}\right)\left(p^{(k)} - H_k q^{(k)}\right)^T}{q^{(k)T}\left(p^{(k)} - H_k q^{(k)}\right)}. \tag{21}$$

Then, we define that

$$\Delta H_k = \frac{p^{(k)} p^{(k)T}}{p^{(k)T} q^{(k)}} - \frac{H_k q^{(k)} q^{(k)T} H_k}{q^{(k)T} H_k q^{(k)}}, \tag{22}$$

$$H_{k+1} = H_k + \frac{p^{(k)} p^{(k)T}}{p^{(k)T} q^{(k)}} - \frac{H_k q^{(k)} q^{(k)T} H_k}{q^{(k)T} H_k q^{(k)}}. \tag{23}$$

We swap H equals B, p and q; therefore,

$$q^{(k)} = B_{k+1} p^{(k)}, \qquad (24)$$

$$B_{k+1} = B_k + \frac{q^{(k)} q^{(k)T}}{q^{(k)T} p^{(k)}} - \frac{B_k p^{(k)} p^{(k)T} B_k}{p^{(k)T} B_k p^{(k)}}. \qquad (25)$$

We assume that $B_{(K+1)}$ is reversible, then,

$$H_{k+1} = B_{k+1}^{-1}. \qquad (26)$$

Finally, the BFGS formula for H was obtained as

$$H_{k+1}^{BFGS} = H_k + \left(1 + \frac{q^{(k)T} H_k q^{(k)}}{p^{(k)T} q^{(k)}}\right) \frac{p^{(k)} p^{(k)T}}{p^{(k)T} q^{(k)}} - \frac{p^{(k)} q^{(k)T} H_k + H_k q^{(k)} p^{(k)T}}{p^{(k)T} q^{(k)}}. \qquad (27)$$

The iteration of the D-matrix can be realized through iterative calculation. Furthermore, storage space can be effectively saved by transforming the matrix into a vector. Therefore, the algorithm convergence process for big data can be realized by another approximation of the BFGS algorithm, which is also called the L-BFGS algorithm.

3.3. Validation of the Model.

The accuracy of the model can be effectively verified through the accuracy test. Accuracy, classification error rate, precision, recall, and $F1\_score$ were the five indicators of score used to evaluate the effectiveness of amchine learning algorithms. The specific formulas are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \qquad (28)$$

$$Classification\,error\,rate = 1 - accuracy, \qquad (29)$$

$$Preciion(p) = \frac{TP}{(TP + FP)}, \qquad (30)$$

$$Recall(R) = \frac{TP}{(TP + FN)}, \qquad (31)$$

$$F1score = \frac{2 * (P * R)}{(P + R)}, \qquad (32)$$

where TP means true positive; TN means true negative; FP means false positive; FN means false negative. In the $F1\_$ score calculation formula, precision is abbreviated as P, and recall is abbreviated as R. The $F1\_score$ value ranges from 0 to 1, with 1 indicating the best and 0 the worst.

4. Results

4.1. Big Data Platform Construction Framework.

The big data infrastructure is a stack-type technology architecture [40], Mainly including the following: (1) the foundation layer; (2) the management level: not only the storage and management of existing data, but also the calculation of some data; (3) the analysis layer: embed the corresponding statistical model and machine learning algorithm to analyze the data according to the research objectives; (4) application layer: mainly for the user's front-end development and visual output, as shown in Figure 5.

4.2. Storage Engine.

As the bottom layer of the big data architecture of disease burden, the analysis layer is included in the platform building framework of the basic layer in this study. The big data platform building ideas of Internet enterprises are used for reference and optimized to form a highly automated computing platform that can be expanded horizontally. Specifically, the Kudu storage engine is used for large-scale data storage to balance the performance of HDFS and HBase random reading and writing and batch analysis [41]. The access to a remote MySQL database is implemented by Faderated, and on this basis, the basic layer of the big data platform for estimating disease burden is built, as shown in Figure 6.

4.3. Big Data Operation.

After the storage engine is selected, since the traditional Hadoop ecosystem cannot complete the ETL and data cleaning work in one MapReduce, the data calculation and processing are realized through the hybrid construction of Spark and MapReduce, as shown in Figure 7.

In order to improve the calculation efficiency and reduce the fault tolerance of the distributed system, the Spark process builds an elastic distributed data set by referring to the idea of functional programming. As a read-only and partitioned data set, RDD forms a directed acyclic graph through operator connection, which significantly improves the computational efficiency. The transformation between various operators is realized through stream, as shown in Figure 8.

Therefore, Spark, MapReduce, and Sparkstreaming jointly complete the cloud computing process of the disease burden big data platform. The sorting and coordination of data resources are jointly completed by Zookeeper and Apache Hadoop Yarn. The zookeeper server processes the client's request through FIFO, allowing the client to connect to any subserver and providing higher performance.

4.4. Multisource Data Acquisition, Cleaning, and Integration.

For the construction of the big data platform for disease burden, the management level should unify the management and identification of structured data and unstructured data. Meanwhile, the real-time data should have the ability for rapid query and error identification, and the system response time should be shortened as much as possible. At the same time, sufficient operable space should be provided for the future upgrading of the system.

Therefore, the data types are mainly divided into three categories, mainly including the following:

(1) Front-end logs: big data from the Internet, medical and health institutions, and mobile phones;

(2) Back-end log: summarize and transmit data from subservers around the world;

(3) data from MySQL database of public security, civil affairs and other institutions.
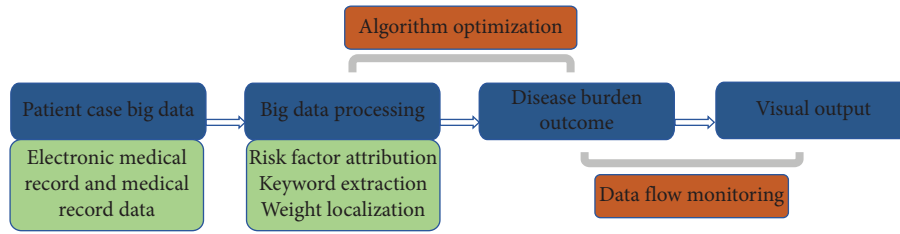
FIGURE 5: Disease burden big data cloud platform architecture.
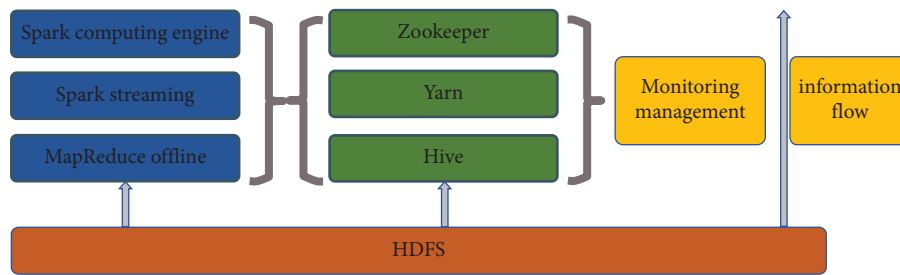
FIGURE 6: Organizational framework and construction idea of the foundation layer.
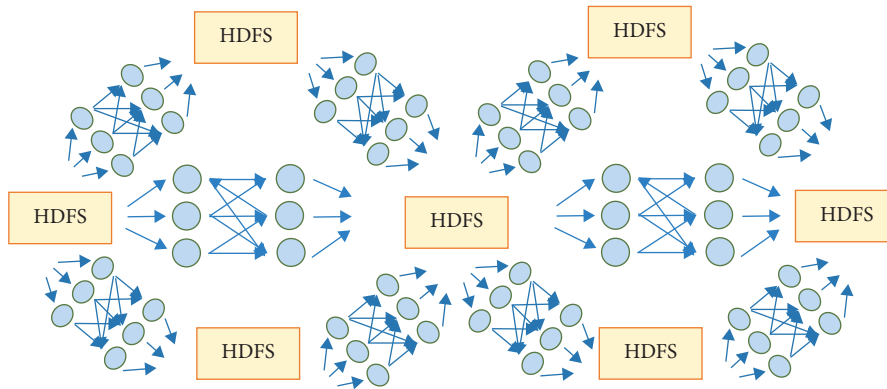
FIGURE 7: MapReduce data processing process.
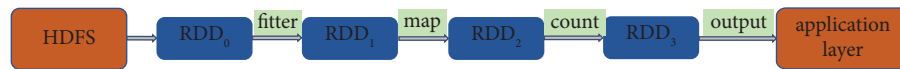
FIGURE 8: Elastic distributed data set framework.

For different types of data flows, the Kafka producer protocol is implemented based on Lua to achieve efficient data collection. The specific framework is shown in Figure 9.

4.5. Call of Log Data. For the front-end log collection and access, it is required to have high reliability and availability while responding in real time [42]. By referring to the

construction ideas and failure cases of Internet enterprises, a large number of tests have compared Flume, Scribe and Chukwa's various construction ideas and frameworks, which cannot meet the collection and storage of super-large amounts of disease burden data. Therefore, a set of data acquisition gateway can be developed based on Kafka to complete data acquisition and realize through nginx Lua. The back-end log collection and access can use Go language
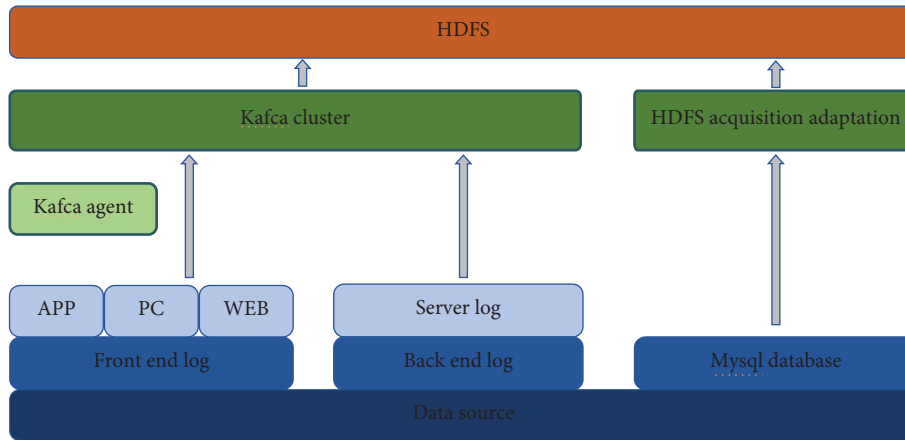
FIGURE 9: Data acquisition framework.
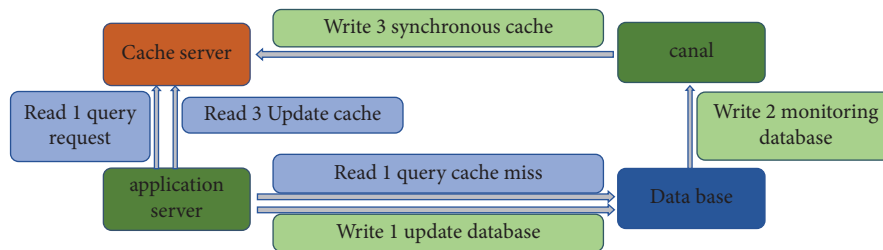
FIGURE 10: MySQL database data collection optimization strategy.

to realize file collection because the server logs are relatively stable. At the same time, multiple reconfigurations and optimizations can be carried out according to the needs of future research. The traditional database collection and access method uses canal to update the cache, which leads to slow MySQL query speed and no QPS. It is easy to cause paralysis if a large number of requests are sent to MySQL. Therefore, we proposed the solution of adding a cache in front of MySQL during the construction, as shown in Figure 10.

Specifically, when the cache is exhausted, MySQL will write another copy to the cache. When the data is inconsistent (MySQL database changes), modify it asynchronously, and then start a canal service to monitor MySQL to make the synchronous cache consistent.

4.6. Application Layer Construction.
In the construction of the application layer, we mainly highlight two functions. The first is the real-time monitoring of various types of data streams on the Internet and the stability monitoring of front-end log data streams. The second is the visual presentation of disease burden measurement results. By using AI to design

the UI interface for the visual output of disease burden results, we hope to improve the big data analysis ability of disease burden in China. The elk monitoring system has the following advantages:

(i) Business data analysis: collect key steps from the front-end information to the back-end for business process analysis.

(ii) Error log analysis: similar to bugly, after the error log is reported, errors can be summarized, displayed by category, and analyzed in the back end.

(iii) Data early warning: with elk, it is easy to establish an early warning mechanism for monitoring fields and conduct early warning before large-scale errors occur.

4.7. Visual Output of Disease Burden Results.
The results of the analysis of disease burden are visually output through the UI interface. Specifically, it includes four modules: an overall overview, data analysis, data prediction, and data application (Figures 11 and 12). This study designed a visual display of a big data platform based on the calculation
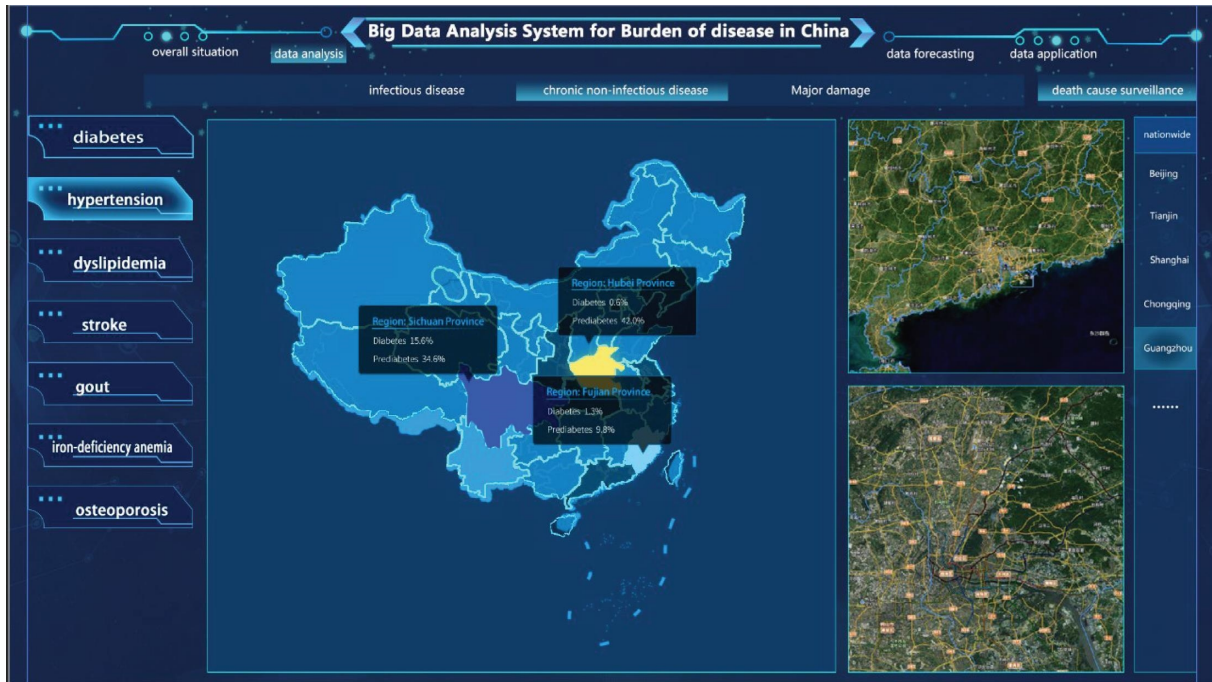
FIGURE 11: Overall overview of China's disease burden big data analysis system.

FIGURE 12: Data analysis module of China disease burden big data analysis system.

results of the burden of diabetes in Guangzhou, China. In the overall overview module, Baidu Map is embedded in the whole disease analysis big data platform, which can realize real-time tracking and analysis of data and visually output the overall situation of diseases according to different disease classifications.

In the data analysis module, the disease burden is mainly measured and demonstrated by big data. Based on the research on regional disease burden, a big data analysis system for urban disease burden has been formed. We try to form an integrated solution from data collection to analysis to decision-making.

5. Conclusion

Medical and health data, especially the massive microdata in the measurement of disease burden, have the characteristics of complex, multisource, and diverse data. At the same time, it also has a complexity and diversity different from other data types, which often makes medical workers and health managers unable to use conventional software tools to acquire, manage, and integrate medical and health data in a short time, making it valuable information. It is particularly important to use big data technology to solve this problem. The construction of the big data platform for disease burden is

different from that of enterprises and businesses. Since the service objects of the big data platform are mainly health department decision-makers and relevant researchers, what application layer can display the analysis results most intuitively and objectively? Can the disease burden analysis results on the IHME website be further improved to enhance readability? Domestic direct research on the construction of a health big data platform is relatively rare, and more research focuses on theoretical research and empirical measurement.

A large number of big data enterprise-level application practices have proved that the disease burden big data platform based on Spark engine can effectively realize the collection and intelligent management of multisource heterogeneous medical data. On the basis of the traditional distributed computing framework, it has greatly optimized and improved different links and levels, especially in the distributed computing of big data, which has high application value and practical significance in this field. In the future, we can apply big data technology to build a unified information management platform and strengthen the construction of data set standards, technical standards, and data sharing and exchange standards, so as to realize the effective application of medical and health data information and promote data integration and information sharing.

There are some shortcomings in this study. (1) Due to the updating of technology, the convergence and quasi-combination of the algorithm do not reach the optimal solution. At the same time, the degree of matching of different data types to the algorithm is not nearly the same, which needs more detailed research in the future. (2) Some prediction algorithms are still under active design and development and are not included in the design framework of the current big data platform. (3) Due to the heavy programming workload required for platform construction, the big data analysis platform built in this study is still under active construction and has not been put into use. The platform will be built and put into use in the future.

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] Z. Zhao and Q. Hu, "The application of a computer monitoring system using IoT technology," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–11, 2022.

[2] G. H. Kim, C. M. Jun, H. C. Jung, and J. Ho Yoon, "Providing service model based on concept and requirements of spatial big data," *Journal of Korean Society for Geospatial Information System*, vol. 24, no. 4, pp. 89–96, 2016.

[3] J. Wang, C. Zeng, Z. Wang, and K. Jiang, "An improved smart key frame extraction algorithm for vehicle target recognition," *Computers & Electrical Engineering*, vol. 97, Article ID 107540, 2022.

[4] C. Li, C. Liao, X. Meng et al., "Effective analysis of inpatient satisfaction: the random forest algorithm," *Patient Preference and Adherence*, vol. 15, pp. 691–703, 2021.

[5] R. Martínez-Castaño, J. C. Pichel, and D. E. Losada, "A big data platform for real time analysis of signs of depression in social media," *International Journal of Environmental Research and Public Health*, vol. 17, no. 13, p. 4752, 2020.

[6] S. Asadianfam, M. Shamsi, and A. Rasouli Kenari, "Big data platform of traffic violation detection system: identifying the risky behaviors of vehicle drivers," *Multimedia Tools and Applications*, vol. 79, no. 33-34, pp. 24645–24684, 2020.

[7] A. A. Munshi and A. Alhindi, "Big data platform for educational analytics," *IEEE Access*, vol. 9, pp. 52883–52890, 2021.

[8] A. A. Munshi and Y. A. R. I. Mohamed, "Data Lake lambda architecture for smart grids big data analytics," *IEEE Access*, vol. 6, pp. 40463–40471, 2018.

[9] P. Tampakis, E. Chondrodima, A. Tritsarolis et al., "i4sea: a big data platform for sea area monitoring and analysis of fishing vessels activity," *Geo-Spatial Information Science*, vol. 25, no. 2, pp. 132–154, 2022.

[10] C. Xu, X. Du, Z. Yan, and X. Fan, "ScienceEarth: a big data platform for remote sensing data processing," *Remote Sensing*, vol. 12, no. 4, p. 607, 2020.

[11] P. E. Bourne, V. Bonazzi, M. Dunn et al., "The NIH big data to knowledge (BD2K) initiative," *Journal of the American Medical Informatics Association*, vol. 22, no. 6, p. 1114, 2015.

[12] P. Vicini, O. Fields, E. Lai et al., "Precision medicine in the age of big data: the present and future role of large-scale unbiased sequencing in drug discovery and development," *Clinical Pharmacology & Therapeutics*, vol. 99, no. 2, pp. 198–207, 2016.

[13] A. G. Vaithinathan and V. Asokan, "Public health and precision medicine share a goal," *Journal of Evidence-based Medicine*, vol. 10, no. 2, pp. 76–80, 2017.

[14] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge," *Contemporary Oncology*, vol. 19, no. 1, pp. A68–A77, 2015.

[15] G. S. Ginsburg and K. A. Phillips, "Precision medicine: from science to value," *Health Affairs*, vol. 37, no. 5, pp. 694–701, 2018.

[16] H. S. Cha, J. M. Jung, S. Y. Shin et al., "The korea cancer big data platform (K-cbp) for cancer research," *International Journal of Environmental Research and Public Health*, vol. 16, no. 13, p. 2290, 2019.

[17] H. Khazaei, C. Mcgreger, and M. Eklund, "Toward a Big Data Healthcare Analytics System: A Mathematical Modeling Perspective," in *Proceedings of the 2014 IEEE World Congress on Services*, 27 June 2014 - 02 July 2014.

[18] http://www.gov.cn/zhengce/content/2016-06/24/content_5085091.html.

[19] http://wjw.ah.gov.cn/public/7001/52074961.html.

[20] http://www.gov.cn/xinwen/2021-12/28/content_5664872.htm.

[21] W. Lie, B. Jiang, and W. Zhao, "Obstetric Imaging Diagnostic Platform Based on Cloud Computing Technology Under the Background of Smart Medical Big Data and Deep Learning," *IEEE Access*, vol. 8, pp. 78265–78278, 2020.

[22] Y. Lu, W. Huang, L. Wang, F. Song, Y. Peng, and J. Peng, "Data-enabled Digestive Medicine: A New Big Data Analytics Platform," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 18, no. 3, pp. 922–931, 2019.

[23] C. Jo, "Cost-of-illness studies: concepts, scopes, and methods," *Clinical and Molecular Hepatology*, vol. 20, no. 4, p. 327, 2014.

[24] D. Brecht, "European burden of disease network: strengthening the collaboration.[J]," *The European Journal of Public Health*, vol. 30, no. 1, 2020.

[25] GBD 2019 Risk Factors Collaborators, "Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019," *Lancet (London, England)*, vol. 396, pp. 1223–1249, 2020.

[26] H. Aoki, T. Kitano, and D. Kitagawa, "Disease burden of congenital cytomegalovirus infection in Japan," *Journal of Infection and Chemotherapy*, vol. 27, no. 2, pp. 161–164, 2021.

[27] R. Qiu, M. Hadzikadic, S. Yu, and L. Yao, "Estimating disease burden using Internet data," *Health Informatics Journal*, vol. 25, no. 4, pp. 1863–1877, 2019.

[28] M. P. Jacob, D Thomas Js, R. Bunch Dustin et al., "Health care and precision medicine research: analysis of a scalable data science platform," *Journal of Medical Internet Research*, vol. 21, no. 4, 2019.

[29] J. Neto, A. M. Moreira, and G. Vargas-Solar, "TRANSMUT-SPARK: Transformation Mutation for Apache Spark," *Software Engineering*, 2021.

[30] V. S. Sharma, A. Afthanorhan, N. C. Barwar, S. Singh, and H. Malik, "A dynamic repository approach for small file management with fast access time on Hadoop cluster: hash based extended Hadoop archive," *IEEE Access*, vol. 10, pp. 36856–36867, 2022.

[31] X. Wen and J. Hu, "SWEclat: a frequent itemset mining algorithm over streaming data using Spark Streaming[J]," *The Journal of Supercomputing*, vol. 76, no. 10, 2020.

[32] G. Bello-Orgaz, J. J. Jung, and D. Camacho, "Social big data: Recent achievements and new challenges," *Information Fusion*, vol. 28, pp. 45–59, 2016.

[33] X. Cai, F. Li, P. Li, L. Ju, and Z. Jia, "SLA-aware energy-efficient scheduling scheme for Hadoop YARN," *The Journal of Supercomputing*, vol. 73, no. 8, pp. 3526–3546, 2017.

[34] B Dong, Q. Zheng, F. Tian et al., "Performance models and dynamic characteristics analysis for HDFS write and read operations: a systematic view," *Journal of Systems and Software*, vol. 93, pp. 132–151, 2014.

[35] N. Papadakis, P. Kefalas, and M. Stilianakakis, "A tool for access to relational databases in natural language," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7894–7900, 2011.

[36] A. S. Alammary, "Arabic questions classification using modified TF-IDF," *IEEE Access*, vol. 9, pp. 95109–95122, 2021.

[37] S. Mai, B. Mahmoud, E. G Sally, A. Reham, and E Ali, "An optimized FP-growth algorithm for discovery of association rules," *The Journal of Supercomputing*, vol. 78, no. 4, 2021.

[38] K. Li, S. Martin, C. R. Rojas, S. Chatterjee, and M. Jansson, "Alternating Strategies with Internal ADMM for Low-Rank Matrix reconstruction," *Signal Processing*, vol. 121, pp. 153–159, 2016.

[39] D. Yang, G. Li, and G. Cheng, "On the efficiency of chaos optimization algorithms for global optimization," *Chaos, Solitons & Fractals*, vol. 34, no. 4, pp. 1366–1375, 2007.

[40] N. Bhushana Samyuel and B. A. Shimray, "Securing IoT Device Communication against Network Flow Attacks with Recursive Internetworking Architecture (RINA)," *ICT Express*, vol. 7, no. 1, pp. 110–114, 2020.

[41] M. S. Wiewiórka, D. P. Wysakowicz, M. J. Okoniewski, and T. Gambin, "Benchmarking distributed data warehouse solutions for storing genomic variant information," *Database*, vol. 2017, Article ID bax049, 2017.

[42] A. Boron, "Front-end circuit for energetic signal data acquisition," *Przeglad Elektrotechniczny*, vol. 845, pp. 119–121, 2008.

Hindawi

Research Article

Influencing Factors of Negative Motivation in College Students' English Learning Relying on the Artificial Neural Network Algorithm

Ping Liu ⓘ

Foreign Language School, Hubei Polytechnic University, Hubei, China

Correspondence should be addressed to Ping Liu; 210043@hbpu.edu.cn

College English has received increasing focus as an important part of the education system. However, the continuous development of English instruction has not simultaneously promoted students' positive learning motivation for English courses. The generation and growth of negative motivation have become a common problem among college students. Students' enthusiasm for learning English courses is gradually fading and teachers' teaching value has also become difficult to guarantee, which seriously affects the normal and orderly progress of education and teaching activities. Therefore, it is very important for the healthy development of English teaching to understand and study the affecting elements of negative motivation in English learning of university students and to provide scientific and effective suggestions for teachers and learners to establish a good teaching and learning attitude. Relying on the interpretation of a negative motivation theory, this paper studies various influencing factors by means of the artificial neural network algorithm. The principal component analysis method is introduced to improve the traditional BP algorithm in terms of the frequency of iterations and the length of computation time, which realizes the accurate and efficient analysis of college students' English learning data. The results of the analysis revealed that the comprehensive error of this algorithm in the analysis of influencing factors was in the range of 0.004 to 0.012. Through the calculation of the eigenvalues and cumulative contribution rate of negative motivation influencing factors, it is found that factors such as the curriculum setting, teaching method, and teacher-student relationship have the greatest influence on students' negative motivation in English learning. The eigenvalues were 1.027, 1.319, and 1.422, respectively. The cumulative contribution rate reached 64.57%, 26.11%, and 23.62%, respectively. From this aspect, it is necessary to improve these aspects in order to eliminate the negative motivation of learning.

1. Introduction

With the development of college English instruction, the status and influence of learning motivation in college students' English learning activities are becoming more and more prominent. As a major factor in guiding college students' English learning behavior, learning motives include not only positive motives but also negative motives. In the current stage of English education, the negative motivation of students' English learning has been significantly enhanced and has gradually become a common phenomenon. More and more students are unable to face English learning positively. This is not conducive not only to the

sound growth of English instruction but also to the cultivation of foreign language talents with high comprehensive quality. Therefore, it is very important to study the affecting elements of college students' negative motivation in English learning and understand the generation of negative motivation to enhance students' positive motivation for learning and promote the sustainable development of English teaching. The artificial neural network (ANN) algorithms have received a great deal of attention and exploration over the past few years and have been successfully applied in a wide range of industrial fields, giving ample play to their own unique strengths and application value. For example, it can be seen in medical diagnosis, machinery manufacturing,

construction engineering, finance, insurance, and other industries. By relying on artificial neural network intelligent algorithms, various industries have also ushered in unprecedented development opportunities in the market and achieved considerable development results. In the study of the affecting elements of negative motivation in English learning among university students, it can comprehensively judge and effectively analyze the entire English learning process of students and draw accurate conclusions based on scientific basis, which has a very important practical value to the improvement and promotion of college English instruction.

Regarding the affecting elements of negative motivation in English learning among university students, numerous academics have carried out in-depth studies. Batubara et al. believed that the learning environment is very important for the cultivation of students. They used the Montessori method to explore the factors that affect students' negative motivation in learning [1]. Raczoski et al. used the expectation-value-cost motivation model to study the motivation and influencing factors of students studying abroad courses. They stated that external effort cost and self-efficacy are negative motivation factors [2]. Hasan investigated the changes in learners' motivation to study English as a second language under the background of Indonesian universities and found out the factors that affect the changes in their negative motivation [3]. Anjum et al. showed that there is a negative and low correlation between the influence of the motivation level and students' English academic performance through the Pearson coefficient correlation survey [4]. Soriano et al. selected participants from various programs at Quirino State University, Cabalogis Campus in the Philippines and explored the students' life experiences and the factors that affect their negative motivation to learn English [5]. Yarlagadda et al. used the Bayesian network analysis method to explore the key factors affecting the negative motivation of college students' English learning. They said that the responsibilities of a learner and the properties of a learner had the greatest adverse effect on it [6]. These studies have used different methods to study the influencing factors of negative motivation in English learning, but the conclusions are not inductive and fail to reveal the complexity of each influencing factor. The research studies on the influencing factors of negative motivation require more accurate analysis results from colleges and universities, for which the artificial neural network algorithm is a good choice.

As a major branch of the intelligent algorithm, the artificial neural network algorithm has a very important application value. Ge investigated the affecting elements of basketball damage in sports instruction based on the ANN and analyzed the causes of knee damage in basketball training in general universities [7]. Allahyari E used an artificial neural network model to analyze the influence of different factors, such as gender, age, education level, and place of residence on people's emotional intelligence level [8]. Dar investigated local antitumor effects using ANN modeling and developed a neural network time series model to evaluate it [9]. Palanichamy used artificial neural network

algorithms to detect motor failures. Compared with traditional methods, it can not only increase the new entry detection function but can also improve the accuracy and stability of system performance [10]. Golnaraghi et al. analyzed the significant impact of changes in labor productivity on the economy through artificial neural networks and then proved through experiments that this method is more useful than statistical regression techniques [11]. Hu et al. built a postmodern media communication perspective system by establishing an in-depth research and learning platform for college marketing based on the AI neural network [12]. On the whole, the ANN algorithm has been widely used in the market, but there are few studies that combine it with the elements that impact negative motivation in English studying among university students. Therefore, it is very necessary to use the ANN algorithm to analyze the affecting elements of college students' negative motivation in English learning, which is important to promote the healthy growth of English education.

This paper relies on the ANN algorithm to study various influencing factors of negative motivation. Through research and investigation, it was found that among 146 students, a total of 113 students had negative motivation in English learning, accounting for about 91.1% of the total number. When analyzing the influencing factors of negative motivation, the error range between the analysis value of the traditional BP algorithm and the actual value was 0.016 to 0.275, while the error range of the PCA-BP algorithm improved by principal component analysis was 0.004 to 0.008 in this paper. It showed that the performance of the improved algorithm in this paper is more rational in the analysis of influencing factors. The improved algorithm was used to calculate the eigenvalue and cumulative contribution rate of the influencing factors, and it was found that the eigenvalue of the curriculum setting factor reached 1.027. The cumulative contribution rate was 64.57%, which had the most significant impact on negative learning motivation.

2. Influencing Factors of Negative Motivation in English Learning

2.1. Overview of Negative Motivation in English Learning. Negative motivation is an important content in the study of learning motivation. There is a close relationship between learning motivation and negative motivation; so, it is necessary to understand learning motivation before studying the influencing factors of college students' negative motivation in English learning [13].

The study of learning motivation originally emerged in the field of psychological studies in the 1930s and has since progressively extended into the area of pedagogical studies. Although the study of learning motivation is becoming a century old, a sufficiently precise and widely accepted definition is yet to be developed owing to the inherent intricacies of learning motivation and the diverse areas of study, perspectives, and standpoints of various academics. Scholars have given different definitions of motivation from different perspectives, which to a certain extent allow people to understand the nature of learning motivation more clearly

and comprehensively. In terms of the hierarchy of needs theory, motivation for learning is intrinsic to people's study actions. It is characterized by internal elements and is also influenced by external elements, as shown in Figure 1. Under the combined effect of internal and external elements, study motivation not only can boost, enhance, sustain, and adjust an individual's studying actions but can also reduce an individual's original degree of study motivation due to the influence of some negative factors. Because of the influence of negative factors, the phenomenon that students' original learning motivation level declines is called the generation of negative learning motivation.

The generation of negative learning motivation can be viewed as a passive aspect of motivation, which is a passive manifestation expressed in the learner's study behavior. From a certain point of view, the generation of negative learning motivation can be regarded as a slow dynamic process, that is, the process of students' English learning positive motivation is gradually weakening. There is an inseparable relationship between the weakening or elimination of college students' motivation in the process of English learning and negative motivation. Analyzing the influencing elements of negative motivation is the most important means and method for teachers to improve their own work and students to reflect on their learning effects. For example, teachers should think about which links in teaching activities will affect students' enthusiasm for learning English; students reflect on their own learning methods and states in the learning process.

The factors that affect the occurrence of negative motivation for learning can also be divided into internal reasons and external reasons. External reasons refer to various factors that will negatively affect students, which exert a negative effect on learners' learning behavior through internal factors, with a significant impact on negative motivation, while internal factors are the main reasons for the decrease in learners' motivation, including self-confidence, attitude, and personality traits. It can be seen from numerous studies that those with strong self-confidence in learning and positive attitudes are slightly less affected by external elements, while people with insufficient self-confidence or negative attitudes are more likely to be influenced by adverse external elements. Therefore, it is believed that of the two elements that influence a decline in motivation, the internal elements of the learner are supposed to be the main cause, while the external influencing elements use the internal elements to affect the motivation of the learners and the size of the influence varies from person to person.

2.2. The Artificial Neural Network Algorithm.
In order to better understand the impact of these internal and external elements on learning negative motivation, this paper combines the ANN algorithm to conduct research studies. As a method for simulating the biological neural network, the ANN algorithm is similar to the human nervous system, the main component of which is neurons, as shown in Figure 2. These neurons have a very important role, which are like nerve cells in the human nervous system. Neural
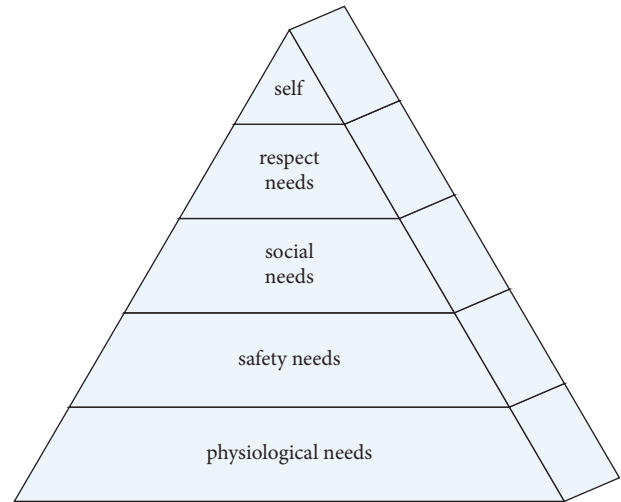
FIGURE 1: The hierarchy of needs.

modules, such as dendrites, axons, and synapses, with different functions are composed of many nerve cells; each of which performs its own function and cooperates with each other. In the analysis of the influencing elements of college students' English learning negative motivation, even if there are factors with high complexity, their function can be realized.

The research on the influencing elements of learning negative motivation is not a simple linear process relying on algorithms but a complex and intersecting processing method. The BP network is a highly nonlinear network, which generally refers to those multilayer forward neural networks that use the error back propagation algorithm, that is, the BP algorithm, and is currently the most widely used neural network in scene problem analyses [14].

In this method, the acquisition, analysis, and identification of the influencing factors of learning negative motivation information are carried out by its algorithm.

Its main components are input layer, hidden layer, and output layer. In addition, it is necessary to use the BP algorithm to continuously adapt and train the weights and thresholds. The selection of the weights and thresholds of the BP network structure has been repeatedly adjusted and trained by the BP algorithm. The topology of the BP network is shown in Figure 3:

In Figure 3, the input value of the student's English learning data is $X_1, X_1, \cdots X_n$ and the output value is $Y_1, Y_1, \cdots Y_m$. The interpretation of each parameter is shown in Table 1.

The BP network function mapping relationship is mainly the mapping from an independent variable to a dependent variable. The specific training steps are divided into 6 steps.

First, network initialization is performed and the threshold sum of the network is initialized.

Second, the output value of the hidden layer is calculated. It is supposed that the output variable is X, and the thresholds and weights in the network input and hidden layers are a and w_{ij}, respectively, which can be expressed as formula (1) [15].
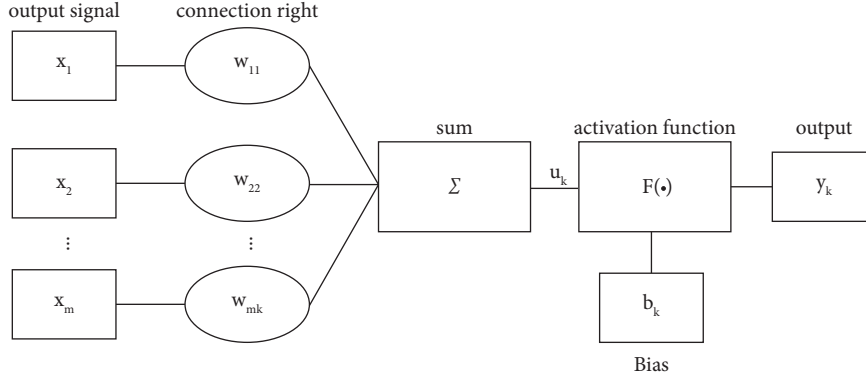
FIGURE 2: Artificial neuron model.
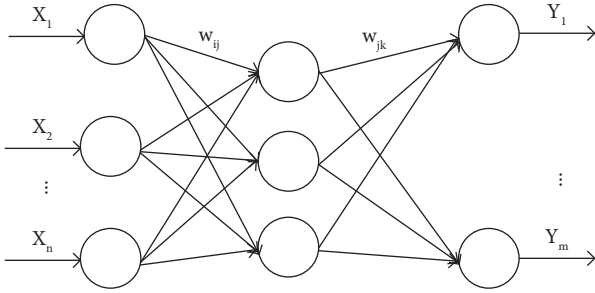
FIGURE 3: BP network topology.

TABLE 1: Interpretation of each parameter.

Sequence	Parameter	Paraphrase
1	w_{ij}	BP network weight
2	w_{jk}	BP network weight
3	n	Network input node
4	m	Network input node

$$H_j = f\left(\sum_{i=1}^{n} w_{ij}x_i - a_i\right), j = 1, 2, \cdots, l. \tag{1}$$

In formula (1), l is the number of hidden layer nodes. f is the activation function of the hidden layer.

Third, the output layer is calculated. H represents the output of the hidden layer. w_{jk} and b represent the connection weight and threshold of the network, respectively. The formula for predicting the output of negative motivation factors is formula (2) [16]:

$$H_k = \sum_{j=1}^{l} H_j w_{jk} - b_k, k = 1, 2, \cdots, m. \tag{2}$$

Fourth, the output error is calculated with reference to the student learning data. Its calculation is mainly carried out through the expected output and predicted output of the negative motivation factors. These two terms are denoted as Y and H, respectively. The error calculation formula is expressed as formula (3) [17].

$$e_k = Y_k - H_k, k = 1, 2, \cdots, m. \tag{3}$$

Fifth, the network input layer and hidden layer weights are updated. The network connection weights w_{ij} and w_{jk} are updated with the prediction error of the network e, as shown in formula (4):

$$w_{ij} = w_{ij} + \eta H_j\left(1 - H_j\right)x(i)\sum_{k=1}^{m} w_{jk}e_k, \tag{4}$$

$$i = 1, 2, \cdots, n \; j = 1, 2, \cdots, m.$$

In formula (4), η is the learning rate.

The sixth step is that the algorithm iteratively judges whether the factors affecting the generation of negative motivation in the English learning data predict the complete end. If it is not over, it will go back to Step 2 until it is over.

However, in this process, the shortcomings of the BP neural network are also very obvious. The first is that the convergence process of the BP neural network model is slow, which leads to the long iterations and computing time required for the training of negative learning motivation factors. The number of layers is difficult to determine. In particular, the number of layers of neurons in the hidden layer of the model and the number of neurons in each layer can only refer to the settings of the empirical structure. This method often causes a lot of computational redundancy in the training of the BP neural network, which affects the normal analysis of learning data.

These problems limit the analysis and research of the BP neural network model on the factors affecting students' negative motivation in English learning to varying degrees. Therefore, this paper uses the PCA method to improve it. PCA is a global algorithm, also known as the principal component analysis method. Its main advantage is that it can provide more information even in the context of less data. It is mainly a basic method to simplify multivariate by linear combination of original variables, as shown in Figure 4.

X is supposed to be the observed student learning data matrix. Each row of it represents an observation and each column represents a variable indicator. These variable indicators will affect the intensity of students' negative motivation to learn at different levels. Among them, there are $X_1, X_2, \cdots, X_p$ variables, with a total of P, and the linear combination of these variables is shown in formula (5) [18].
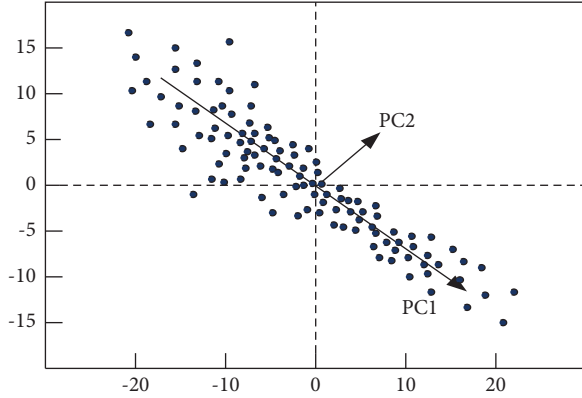
FIGURE 4: Principle of the PCA method.

$$\begin{cases} Y_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\ Y_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\ \cdots\cdots\cdots \\ Y_p = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p \end{cases}. \quad (5)$$

Among them, if P comprehensive indexes $Y_1, Y_2, \cdots, Y_p$ are obtained, the matrix transformation needs to meet the conditions [19].

Unit vectors are all coefficient vectors, as shown in formula (6) [20].

$$a_i' = (a_{i1}, a_{i2}, \cdots, a_{ip}), \quad (6)$$

$a_i'a_i = 1$, which satisfies formula (7):

$$a_{i1}^2 + a_{i2}^2 + \cdots a_{ip}^2, (i = 1, 2, p). \quad (7)$$

There is no relationship between Y_i and $Y_j (i \neq j, i, j = 1, 2, \cdots, p)$.

Y_1 is the variance max in all linear combinations of $X_1, X_2, \cdots, X_p$. Y_2 is the variance max in all linear combinations of $X_1, X_2, \cdots, X_p$ uncorrelated with Y_1. Y_p is the variance max in all linear combinations of $X_1, X_2, \cdots, X_p$ uncorrelated with $Y_1, Y_2, \cdots, Y_{p-1}$.

From these conditions, the first main ingredient, the second main ingredient $\cdots$ the m-th main ingredient in the influencing factor category are obtained in turn.

If the first m principal component is selected to replace the original p independent variables ($m < p$), the dimension of the data will be reduced.

The steps are divided into the following 4 steps:

The first step is to calculate the original data covariance matrix $\sum$. The solution of the eigenvalues of matrix $\sum$ is as formula (8):

$$\lambda_1 \geq \lambda_1 \geq \cdots \geq \lambda_p > 0. \quad (8)$$

The corresponding unit eigenvector is $a_1, a_2, \cdots, a_p$. The variance of the i-th principal component is λ_i, and $i = 1, 2, \cdots, p$.

The next step is to calculate the variance contribution of each main ingredient, as shown in formula (9):

$$\varphi_i = \frac{\lambda_i}{\sum_{j=1}^{P} \lambda_j}, i = 1, 2, \cdots, p. \quad (9)$$

The variance contribution rate of each main ingredient reflects the amount of information. The reflection of the information content of the principal components is weakened in turn. At the same time, the cumulative variance contribution rate of the variance of the first M main ingredients is calculated as formula (10):

$$\Psi_m = \frac{\sum_{i=1}^{m} \lambda_j}{\sum_{j=1}^{P} \lambda_j}, m < p. \quad (10)$$

Among them, when the cumulative variance contribution rate reaches more than 20%, it is considered that this factor has a greater impact on the negative motivation of English learning. If the value reaches more than 60%, it indicates that the influencing factors have a significant influence on negative motivation. At this time, the original p variables can be replaced by m principal components.

Finally, the scores of the observed samples on the principal components are calculated as formula (11):

$$Y_i = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p, i = 1, 2, \cdots, m. \quad (11)$$

If these four steps are satisfied, the principal component analysis can be performed by converting the original dimensional data into the dimension-reduced dimensional data. Similarly, the higher the score is, the more prominent the influence of such factors in the negative motivation of English learning will be.

2.3. Evidence of Influencing Factors.

For the purpose of testing the effectiveness of the research on the affecting elements of college students' negative motivation in English learning relying on the ANN algorithm, this paper investigated the negative motivation of 2021 students in a university. Then, the PCA-BP neural network algorithm was used to analyze the influencing factors of negative motivation on the survey sample data. It was compared with the traditional BP neural network algorithm to verify the feasibility of this method.

Questionnaire survey: This paper conducted a questionnaire survey on the students of the four majors of tourism management, primary education, civil engineering, and computer science in the 2021 grade of the university. The college English course of the school belonged to the compulsory courses of these 4 majors. Students needed to complete 6 credits in 1 academic year to meet the English study requirements. For the purpose of guaranteeing the usefulness of the study, this paper randomly selected 158 students from these four majors as the object of investigation on the influencing factors of negative motivation in English learning. The questionnaire data and details of the experimental subjects are shown in Tables 2 and 3.

It can be seen from Table 2 and Table 3 that the reliability coefficient of the questionnaire in this paper reached 0.862, which met the reliability standard (Generally speaking, if the reliability coefficient of the questionnaire is below 0.6, it

TABLE 2: Details of the experimental questionnaire.

Sequence	Project	Data
1	Number of releases	158
2	Number of recycling	146
3	Efficient	92.4%
4	Number of questions	16
5	Reliability coefficient	0.862

TABLE 3: Details of the subjects of this paper.

Professional category	Number of boys	Number of girls	Total
Tourism management	12	22	34
Elementary education	18	13	31
Civil engineering	18	6	24
Computer science	30	27	57

means that the reliability of the questionnaire needs to be verified. If the reliability coefficient is above 0.8, it means that the reliability of the questionnaire is good). The total number of boys selected in the experiment was 78 and the total number of girls was 68. The ratio of male to female was also relatively balanced.

The survey of English learning negative motivation of students of various majors is shown in Figure 5.

Figure 5(a) shows the number of people who did not generate negative motivation. Figure 5(b) shows the number of people with negative motivation.

It can be seen from Figure 5 that among the subjects of each major, the number of people who have negative motivation for English learning was far more than the number of people who did not have negative motivation. According to the questionnaire data, among the 146 students, a total of 113 students had negative motivation in English learning, accounting for about 91.1% of the total number. Among them, there were 26 people majoring in tourism management, accounting for 76.5% of the total number. There were 27 students majoring in primary education, accounting for 79.4% of the total. There were 21 people majoring in civil engineering, accounting for 91.3% of the total number. There were 39 people majoring in computer science, accounting for 68.4% of the total number. In general, most students were seriously affected by negative motivation in English learning, which hindered the growth of learning enthusiasm and disrupted the orderliness of learning activities. This kind of performance was more obvious in the civil engineering profession, which might be related to the professional characteristics. The gender ratio of this major was heavily biased towards boys, so the trend of negative motivation in English learning was more obvious, which also proved the necessity of the experiment in this paper.

Data analysis: The survey data of these 113 students with negative motivation to learn were used as an independent sample for the algorithm analysis. The BP algorithm and the modified PCA-BP algorithm in this paper were used to analyze and study the influencing factors of negative motivation. In the training process of the two algorithms, the training set was used as the input of the neural network. The target error of the network was set as goal = 0.01, and the

training times were set as 10,000 times. The learning rate was set to 0.75. The sigmoid tangent function tansig was selected as the neuron function of the hidden level and the output level in the network and the training function of the network was training. The MATLAB software (2022a version) was used to establish an analysis model for the influencing factors of negative motivation in English learning, and the results are shown in Figure 6.

Figure 6(a) shows the analysis of the BP algorithm. Figure 6(b) shows the analysis of the PCA-BP algorithm.

According to the survey items, the influencing factors of students' negative motivation in English learning are divided into 7 categories, namely, teaching equipment factors, curriculum setting factors, teaching method factors, teaching material factors, teacher-student relationship factors, learning difficulties factors, and personality factors. As can be seen from Figure 6, the traditional BP neural network algorithm was not very satisfactory in the overall analysis of the negative motivation factors in the survey data. The first four items, as external factors, are of objective nature, and the complexity of the influencing factors is not high. The error range between the analytical results and the actual results of the traditional BP algorithm was 0.016 to 0.275. The error range of the modified algorithm in this paper was 0.004 to 0.008. The latter three are internal factors. In the analysis of the affecting elements of negative motivation, their correlation is wider and the complexity is higher, and they vary due to individual differences of students. The error range between the analytical value and the actual value of the traditional BP algorithm was 0.019 to 0.304. The error range of the improved algorithm in this paper was 0.008 to 0.012.

In the process of students' learning, due to the setting or changes of some external conditions, it is very likely that some students' English learning motivation will be weakened. From the survey data of this paper, if the English teaching equipment, teaching materials, curriculum, and teaching methods are not planned scientifically and systematically, it will easily lead to the weakening of students' motivation. However, internal factors such as the teacher-student relationship, learning difficulties, and students' personalities have no weaker effect on negative motivation than external factors, which are often more complex. Without an effective analysis of the relationship between these factors, it is impossible to provide scientific decision-making to improve the negative status quo of English teaching and learning. From the analysis error, it can be seen that the improved PCA-BP algorithm in this paper can analyze the influencing factors more accurately.

In order to further understand the specific impact of these factors on negative motivation, this paper uses the PCA-BP algorithm to calculate the eigenvalues and cumulative contribution rates of various factors in the learning process, as shown in Figure 7.

Figure 7(a) shows the characteristic values of the influencing factors. Figure 7(b) shows the cumulative contribution rate of influencing factors.

Generally speaking, an eigenvalue greater than 1 indicates that the factor selection is reasonable. The greater the cumulative contribution rate is, the greater the effect of the

(a)

(b)
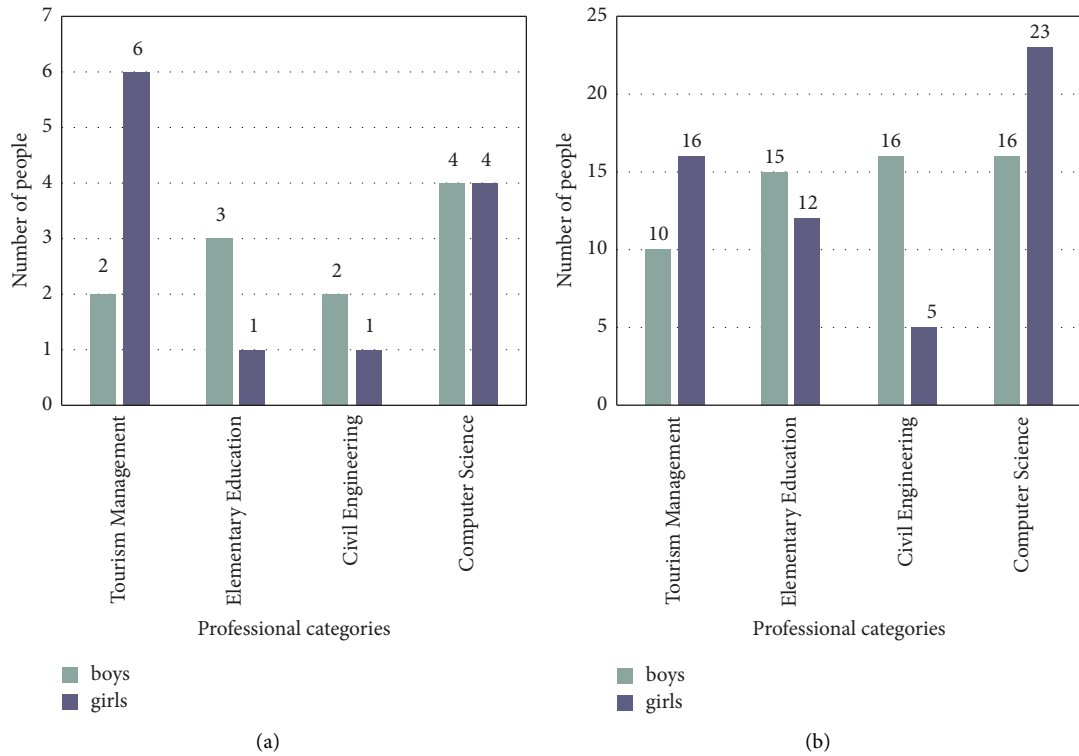
Figure 5: A survey of negative motivation in English learning.

(a)

(b)
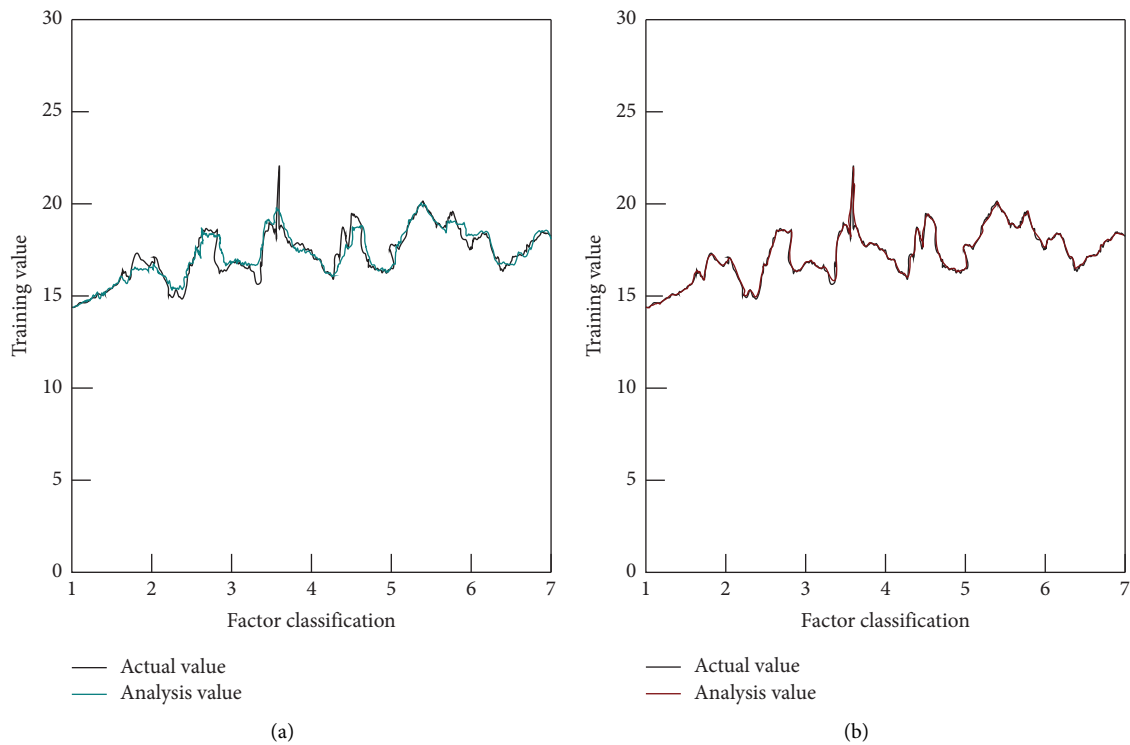
Figure 6: Analysis of influencing factors of negative motivation.

influencing factor will be. It is evident from Figure 7 that the characteristic values of the seven types of factors were 4.364, 1.027, 1.319, 1.571, 1.422, 3.281, and 1.617, respectively, which were all greater than 1. It showed that in the experiment of this paper, the factors screened by the survey have a significant impact on negative motivation. The cumulative contribution rates are 11.34%, 64.75%, 26.11%, 17.29%, 23.62%, 12.35%, and 18.38%, respectively. Among
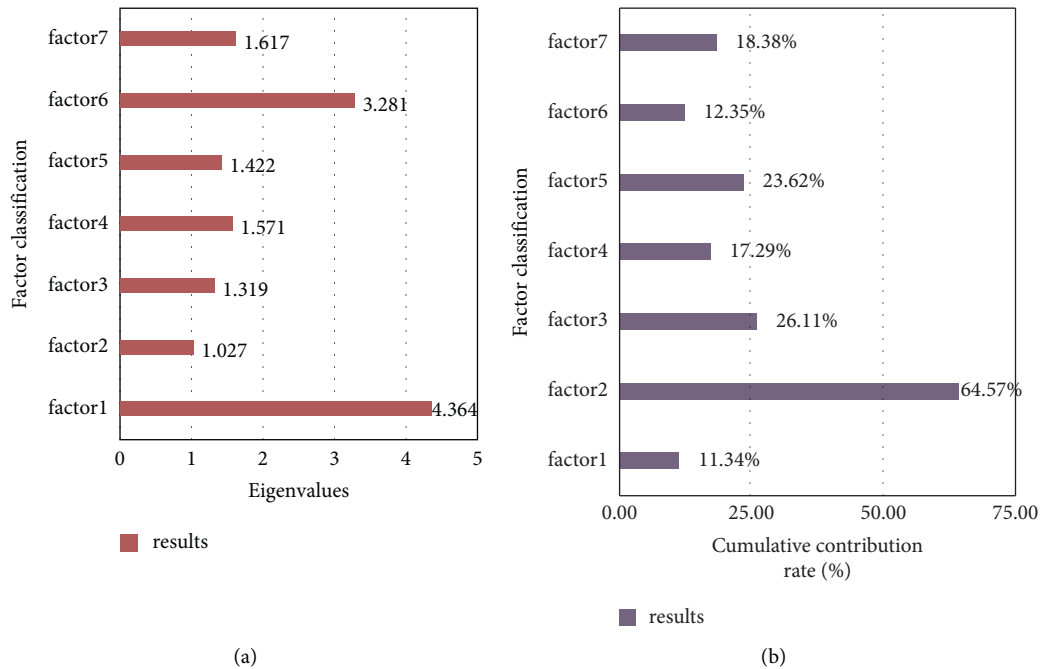
FIGURE 7: Eigenvalues and cumulative contribution rates of various factors.

them, the cumulative contribution rate of curriculum setting is the largest, followed by teaching methods, teacher-student relationship, student personality, learning difficulties, and teaching equipment.

According to the data analysis, the curriculum setting elements have the most significant impact on the negative motivation of students' English learning, which was much higher than that of other factors. The reason why the curriculum setting factor has become the most influential factor is mainly related to the large reduction of college English class hours in recent years and the nature of college English courses. This greatly exacerbates the contradiction between the time-consuming and labor-intensive characteristics of college English learning and the serious shortage of teaching hours. The mechanical repetition of teaching form and content will inevitably lead to students losing interest in learning, making them feel tired of learning and creating negative motivation to learn English. The two influencing factors of teaching style and teacher-student relationship also have a great influence on negative motivation. A single boring teaching method and a negative teacher-student relationship will make students' enthusiasm for English learning subside. Although students' personality, teaching methods, learning difficulties, and teaching equipment have less impact on negative motivation than the first three items, they also have important reference value in improving the English teaching and learning process.

3. Conclusion

Among college students, as a relatively common negative phenomenon, the occurrence of negative motivation in English learning poses a great challenge to the value and significance of English education and teaching. Combined with the ANN algorithm, this paper examines the influencing factors of students' negative motivation in learning and finds that factors such as curriculum setting, teaching methods, and teacher-student relationship have a great influence and correlation with the generation of negative motivation. If the curriculum and teaching are not organized scientifically and effectively or a good and a harmonious relationship between instructors and learners is not formed, the chances of students losing their enthusiasm and motivation for English courses are greater. Therefore, for the purpose of ameliorating these problems, it is important to combine the teaching practice from the perspective of these influencing factors to raise the instruction level and learning quality, so as to eliminate the negative motivation of students. Due to some objective reasons, there are still several areas that need to be improved in this paper, and the experimental level of this paper is also limited. Only one university was selected as the sample, and more differentiated and diversified sample data were not introduced, which made the conclusions of this paper still have certain limitations. In future research, algorithms will be combined to explore the influencing factors of negative motivation from different perspectives and scopes.

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] F. Batubara, T. Derin, N. S. Putri, and R. S. Yudar, "Five factors influencing the students' motivation to learn English as a foreign language: a closer look into Montessori classroom environment," *REiLA Journal of Research and Innovation in Language*, vol. 2, no. 2, pp. 76–84, 2020.

[2] B. Raczoski, J. S. Robinson, M. C. Edwards, and M. Baker, "Forecasting college students' motivations to study abroad: a pilot study," *Journal of Agricultural Education*, vol. 59, no. 2, pp. 123–142, 2018.

[3] D. C. Hasan, "What factors influence changes in students' motivation to learn English as a foreign language?" *Asian EFL Journal*, vol. 20, no. 2, pp. 20–32, 2018.

[4] S. Anjum, F. L. Naz, and A. Quddus, "Analysis of impact on mental health and motivation levels of higher-levels students due to novel virus COVID-19," *Sjesr*, vol. 4, no. 2, pp. 431–437, 2021.

[5] R. M. Soriano, C. N. Escario, P. B. Cagurangan et al., "The role of motivation in English language learning: a qualitative study," *Asia Proceedings of Social Sciences*, vol. 8, no. 2, pp. 68–72, 2021.

[6] P. K. D. V. Yarlagadda, J. Sharma, and P. Silva, "Factors influencing the success of culturally and linguistically diverse students in engineering and information technology," *International Journal of Engineering Education*, vol. 34, no. 4, pp. 1384–1399, 2018.

[7] S. Ge, "Research on the factors of basketball injury in physical teaching based on artificial neural network," *Revista de la Facultad de Ingenieria*, vol. 32, no. 3, pp. 415–422, 2017.

[8] E. Allahyari, "Application of artificial neural network in predicting EI," *Biomedicine*, vol. 10, no. 3, pp. 18–24, 2020.

[9] M. S. Dar, K. B. Akram, A. Sohail et al., "Heat induction in two-dimensional graphene–Fe3O4 nanohybrids for magnetic hyperthermia applications with artificial neural network modeling," *RSC Advances*, vol. 11, no. 35, Article ID 21702, 2021.

[10] P. Palanichamy, "Application of artificial neural network in electrical power system," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 9, no. 1, pp. 77–80, 2018.

[11] S. Golnaraghi, Z. Zangenehmadar, O. Moselhi, and S. Alkass, "Application of artificial neural network(s) in predicting formwork labour productivity," *Advances in Civil Engineering*, vol. 2019, Article ID 5972620, 11 pages, 2019.

[12] J. Hu, B. Liu, and H. Peng, "Role of AI for application of marketing teaching –A research perspective1," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 1–9, 2020.

[13] C. H. Tsai, C. H. Cheng, D. Y. Yeh, and S. Y. Lin, "Can learning motivation predict learning achievement? A case study of a mobile game-based English learning approach," *Education and Information Technologies*, vol. 22, no. 5, pp. 2159–2173, 2017.

[14] S. Beirovi, "The relationship between gender, motivation and achievement in learning English as a foreign language," *European Journal of Contemporary Education*, vol. 6, no. 2, pp. 210–220, 2017.

[15] D. Xing and B. Bolden, "Exploring oral English learning motivation in Chinese international students with low oral English proficiency," *Journal of International Students*, vol. 9, no. 3, pp. 834–855, 2019.

[16] K. H. Rha, "Analyzing the impact of the EFL college students' English learning motivation and strategies on their achievement," *Foreign Languages Education*, vol. 26, no. 1, pp. 91–107, 2019.

[17] E. Alieto and J. Torres, "English learning motivation and self-efficacy of Filipino senior high school students," *Asian EFL Journal*, vol. 22, no. 1, pp. 51–72, 2019.

[18] K. J. Kim, "Changes in English learning motivation of high school students: motivation, demotivation, and remotivation," *English teaching*, vol. 74, no. 4, pp. 249–270, 2019.

[19] U. Zuperl, F. Cus, A. Zawada-Tomkiewicz, and K. Stepien, "Neuro-mechanistic model for cutting force prediction in helical end milling of metal materials layered in multiple directions," *Advances in Production Engineering & Management*, vol. 15, no. 1, pp. 5–17, 2020.

[20] D. Wang, K. Tan, Y. Dong, G. Yuan, and X. Du, "Estimating the position and orientation of a mobile robot using neural network framework based on combined square-root cubature Kalman filter and simultaneous localization and mapping," *Advances in Production Engineering & Management*, vol. 15, no. 1, pp. 31–43, 2020.

Research Article

A Study on THE Mechanism of Electroacupuncture to Alleviate Visceral Pain and NGF Expression

Dongdong Liang ⓘ, Yelong Ren ⓘ, Ledan Huang ⓘ, and Shenhui Jin ⓘ

Department of Anesthesiology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, Zhejiang, China

Correspondence should be addressed to Shenhui Jin; jinshenhui@mjc-edu.cn

Visceral pain is unbearable, and natural methods are needed to relieve it. Electroacupuncture is a relatively new technique that helps relieve visceral pain by improving blood circulation and providing energy to clogged parts of the body. However, its analgesic effect and mechanism in colorectal pain are still unknown. In this study, the visceral pain models of electroacupuncture in rats were compared and discussed, using nanocomponents to stimulate the expression and mechanism of the nerve growth factor in colorectal pain and electroacupuncture and to observe the expression and mechanism of nerve growth factor in visceral pain relief rats induced by nanocomponents and electroacupuncture. The results show that nanocomponents can effectively relieve visceral pain under the action of electroacupuncture. NGF can activate endogenous proliferation, migration, differentiation, and integration. NSC can promote nerve regeneration and recovery after injury.

1. Introduction

Electroacupuncture can be used for long-term stimulation of one or more waveforms of the same intensity or varying intensity [1]. Therefore, it also has the function of traditional acupuncture and moxibustion, namely, to dig the channels and collaterals and coordinate, and it also has the electric current stimulation effect [2, 3]. It can act on some painful parts of the body surface and has an immediate analgesic effect, i.e., long time, poor analgesic effect. The spare wave can cause muscle contraction, produce a strong sense of tremor, improve muscle ligament tension, regulate vasomotor function, improve blood circulation, promote recovery of neuromuscular function, and have a good therapeutic effect on neuromuscular paralysis [4, 5]. The density wave is a combination of density wave and density wave. The duration of density is about 1.5 s. It can produce immediate and delayed inhibition in sensory and motor nerves, play a strong role in analgesia and maintenance, and the tissue is not able to easily produce an adaptive response. It is commonly used in acupuncture analgesia. By adjusting

the frequency, continuous waves become a single sparse wave or dense wave [6, 7]. In the form of dense wave, the pain was relieved immediately, and in the form of sparse wave, the pain relief effect was maintained again, and the wave width is the same, and the patient is not uncomfortable. The intermittent wave can improve muscle excitability, but it has no obvious analgesic effect. Irritable bowel syndrome (IBS) is a typical representative of chronic functional visceral pain [8].

After acupoint electroacupuncture treatment, the synthesis and utilization of norepinephrine is accelerated, but the utilization speed is faster than the synthesis speed [9]. The effects of acupuncture on the content of E and NE in plasma are closely related. The effect of acupuncture on a patient with the decrease of the content of E and NE in plasma is better, and the effect on a patient with more release is worse [10, 11]. Low-frequency 2 Hz electroacupuncture can release a lot of enkephalin in the spinal cord, and high-frequency 100 Hz electroacupuncture can release a lot of enkephalin in the spinal cord [12]. The 2–100 Hz density wave can cause the release of enkephalin and myorphine

simultaneously. Enkephalin and muscle enkephalin enhance each other and have a strong analgesic effect on pain caused by visceral stimulation [13].

In the central nervous system, there are not only pain centers with nociceptive stimulation but also downward regulating systems, which deal with various stimulation information and are related to analgesia [14]. As we all know, the stimulation of entering the spinal cord pain is transmitted through the peripheral nerve, and then from the brainstem and thalamus, limbic system and cerebral cortex on the new and old spinal cord tract, and other stimulation information integration and processing at all levels of the center, finally produce the emotional response of pain and pain [15]. Acupuncture is believed to not only inhibit the transmission of pain information from the spinal cord to the central system but also inhibit pain and pain response by stimulating the human body's neurohumoral and pain-regulating system so as to play an analgesic role [16]. However, its analgesic effect and mechanism in colorectal pain are still unknown.

In this study, the visceral pain models of electro-acupuncture in rats were compared and discussed, using nanocomponents to stimulate the expression and mechanism of nerve growth factor in colorectal pain and electroacupuncture and to observe the expression and mechanism of nerve growth factor in visceral pain relief rats induced by nanocomponents and electroacupuncture. The results show that nanocomponents can effectively relieve visceral pain under the action of electroacupuncture. NGF can activate endogenous proliferation, migration, differentiation, and integration. NSC can promote nerve regeneration and nerve function recovery after injury.

2. Literature Review

2.1. Visceral Pain in Rats. There are peripheral and central sensitization mechanisms in visceral hypersensitivity, but most previous studies are limited to the peripheral sensitization mechanism and lack of understanding of the role of spinal cord, especially the central level above the spinal cord. Neurotransmitters 5-hydroxytryptamine (5-HT), calcitonin gene-related peptide (CGRP), and nitric oxide in the visceral sensory pathway play an important role in the spinal cord and above [17, 18]. It can be seen that the formation of visceral sensation is a complex process, and the viscera, spinal cord, and above ones participate in the perception of visceral stimulation. Acupuncture is an effective way to treat pain. However, as the response point and stimulation point of acupuncture diagnosis and treatment, the essence and mechanism of acupuncture point are not clear.

"Zusanli" and "Shangjuxu" are commonly used acupoints in clinical treatment of gastrointestinal diseases [19]. Acupuncture can stimulate a variety of receptors, produce acupuncture signals, and reach the spinal cord, medulla oblongata, thalamus, and cerebral cortex through different ways [20]. The spinal cord is the first station to deal with and translate acupuncture analgesia; the brainstem is the relay station of forepart, syndrome differentiation, stimulation, synthesis, and transfixion, which plays an important role in

acupuncture analgesia; the thalamus part is the coordination center to strengthen acupuncture analgesia and control analgesia, with a variety of neurohumoral system involved in complex analysis and comprehensive regulation of various information; the limbic system, its nucleus, and various neurotransmitters play a coordinating role in acupuncture analgesia; and the cerebral cortex is the highest center, which is not only a process of excitation and inhibition but also a complex center of regulation and command, which can not only strengthen analgesia but also inhibits excessively and play a role in maintaining dynamic balance [21, 22]. Electroacupuncture is based on the development of acupuncture and the theory of modern electronic medicine.

The CGRP expression in the trigeminal sensory nucleus, the raphe nucleus, and the solitary tract nucleus of visceral hypersensitive rats increased [23]. CGRP is widely distributed in the sensory nerve fiber endings and tissues of the central and peripheral nervous systems of mammals and humans. The content of CGRP in the central nervous system is the highest in the thalamus and other regions. The functions of different brain regions of CGRP can be completely reversed; the part of the nucleus of neurons expressing CGRP is the place where the visceral primary afferent nerve stops directly, which is related to pain conduction, while a part of the nucleus which inhibits the transmission of the pain nerve and plays an analgesic role, and there is a fiber connection between these nuclei [24, 25].

According to the latest research data, neurotrophic factors have many functions. In the process of nerve regeneration, it can guide the growth of axon, stimulate the growth of axon, and reduce the risk of neuron damage. The external intervention of the above neurotrophic factors shows that neurotrophic factors can promote neuronal regeneration, protect the damage of neurons, and play a certain role in axon regeneration and guidance of axons. In the past, it was thought that the central nervous system could not regenerate after axon injury due to the lack of neurotrophic factors in the microenvironment. Under the promotion of 5-transtryptamine (5-HT), neurons precursor cells began to differentiate into neurons. Moreover, it can promote the differentiation of cortical neurons and hippocampal neurons. Only after cerebral ischemia, the release of 5-HT will increase to meet the needs of the body and participate in the regulation of neurons. It can be seen that there are many injury factors and anti-injury factors in the ischemic brain tissue after cerebral infarction. They coexist and influence each other. Among these factors, some can be intervened by people, some cannot be intervened by people, at least there is no way to intervene at present. There are many factors that can be intervened in, such as the neurotrophic factor, nerve regeneration inhibitory factor, nerve regeneration guiding factor, and so on.

2.2. Electroacupuncture. Electroacupuncture has a good regulatory effect on neurotrophic factors after cerebral ischemia, so we consider whether we can use electroacupuncture to promote the proliferation of NSC in the brain to carry out the treatment of endogenous activation

treatment. After cerebral ischemia, electroacupuncture can effectively stimulate hippocampal neurons, activate the proliferation of hippocampal neurons, and finally restore the neural function of the brain.

At the beginning, traditional Chinese medicine put forward the mechanism of electroacupuncture and described it very clearly. Western medicine is not mature, and the understanding is not in-depth but more just for the metabolism of the body. For example, electroacupuncture can regulate abnormalities of hemorheology, blood lipids, and neurobiochemistry, indirectly improve brain oxygen and energy metabolism or directly expand blood vessels, promote the formation of new blood vessels, and increase local brain oxygen. However, there are few studies on its deep cellular and molecular mechanism.

Electroacupuncture can produce nerve conduction impulses and transmit them in the form of electric energy, which can adjust the excitability of the reached parts and affect the electrical activity of the brain. Like any other tissue and cell regeneration, nerve regeneration includes three stages such as proliferation, migration, and differentiation, which belong to the basic development process of the body. In the stage of cell proliferation, the NSC divides asymmetrically with the proliferation of neural precursor cells. Then neural precursor cells begin to proliferate, migrate, and differentiate. The final result of differentiation is the production of mature neural cells. Adult central nerve regeneration plays a role in the treatment of brain injury and participates in some pathological processes. After brain injury, some factors can stimulate the proliferation and differentiation of neural precursor cells and then migrate to the injured site to play a role instead of the injured cells. Whether it is systemic or focal cerebral ischemia, the central nervous system of the adult brain has this function, which is regulated by the body. In order to stimulate the proliferation and differentiation of precursor cells and form new nerve cells, growth factors were injected. The newly formed neurons have the neuron function. They can also differentiate to meet the needs of the body, receive and send pulses, and then improve the neurological function. When the brain is injured, after a series of complex reactions, it can rapidly stimulate the proliferation and differentiation of neural precursor cells. Therefore, regeneration of the brain nerve that receives signals is the expression ability of the brain, plays a role in its own repair, and plays a role in the treatment process.

After cerebral ischemia, the expression of the NTF gene was activated. In the early stage of ischemia, the NTF gene expression is the most obvious, usually within minutes to hours, but the expression will not last for a long time and will soon return to the original state.

At present, acupuncture analgesia has been used in the treatment of acute and chronic abdominal pain, headache, dysmenorrhea, chronic inflammatory arthralgia, and other sporadic reports, as well as in appendectomy, thyroidectomy, assisted major surgery analgesia, and other related reports. The operation of acupuncture analgesia has some disadvantages, such as lack of pain control, muscle tension, and visceral traction reaction, but the requirement of an auxiliary examination, such as enteroscopy, is relatively low. Acupuncture stimulates the transmission of direct type I and type II afferent nerve fibers to pass through these fibers, and muscles and nerve impulses are transmitted to the anterior and lateral sides of the spinal cord, thereby preventing pain signals from being transmitted up to the spinal tract. Acupuncture stimulates the midbrain-related nuclei by activating the gray matter cells around the aqueduct and the spinal cord nucleus and then transmits signals downward through the dorsolateral tract to reduce the release of norepinephrine and serotonin in the spinal cord. These neurotransmitters inhibit presynaptic and postsynaptic space pain stimulation by inhibiting signal transduction of the spinal thalamic tract. In the pituitary hypothalamic system, acupuncture signals promote enkephalin and adrenocorticotropic hormones to enter the blood.

Acupuncture can also activate the GABA receptor, release the substance P, activate the secretion of vasopressin in the paraventricular nucleus of the hypothalamus, and participate in pain inhibition. The nerve endings and the peripheral autonomic nervous system form a synaptic network in the gastric wall. When the digestive tract is stimulated by the injury, the body can feel and transmit the pain through the activities of internal ENS and pan, and change the intestinal motility and functional structure of the ENS through the intestinal reflex pathway. NOS plays an important role in the transmission and regulation of harmful information in the nervous system. In intestinal motility, NOS is a strong inhibitory transmitter. When the intestinal tract is stimulated by injury, no catalysis-produced blood vessels can relax. NOS can protect the intestinal mucosa by increasing the blood flow of the gastric mucosa.

2.3. NGF Expression Mechanism. Because the adult central nervous system can regenerate and respond to various brain injuries, it has great clinical application value and provides a new treatment basis for the treatment and rehabilitation of brain injury diseases. Its mechanism is to promote the regeneration of adult central nervous system by activating neural precursor cells. Clinically, it is considered that the embryonic central nervous system is the most regenerated, and the mature central nervous system is composed of differentiated and mature nerve cells. Due to external or internal reasons, it is not easy to regenerate after stimulation and repair after injury. Neurotrophic factors play an important role in nerve regeneration.

NGF is the most important bioactive substance in the central nervous system. Increased expression of endogenous NGF in the ischemic brain has a protective effect on neurons. It can resist hypoxia, hypoglycemia, and nitric oxide damage, and effectively protect neurons. Therefore, intraperitoneal injection of exogenous NGF can effectively protect rats and reduce the probability of hypoxic-ischemic brain damage. Neurotrophic factor is a kind of soluble peptide factor that can provide nutrition. It can create a suitable microenvironment for the nervous system and play a role. The condition of producing the neurotrophic function is to combine with the corresponding receptor. In the

Equipment name	Instrument manufacturer
Rat tail cuff blood pressure meter (BP-98A)	Japan Softron corporation
Electric heating constant temperature drying oven (202-1)	Shanghai Lunan scientific instrument factory
Image-Pro Plus 6.0 Image Analysis System	Media cybernetics
Electrotherapy	Kang Ling company

mother, the normal development of the nervous system and the survival of neurons cannot be completed by themselves, which requires a kind of NTF to maintain and can induce the growth direction of the nerve process. After birth, NTF plays an important role in maintaining and regulating various functions and normal physiology of the nervous system.

Neurotrophic factors are not the determinants of axon growth because even if there are neurotrophic factors in the experiment, axons of neurons are still difficult to regenerate, and the peripheral nerve and central nerve can regenerate under the same conditions. These inhibitory proteins exist mainly in the white matter of the central nervous system, inhibit the growth of nerve processes, prevent the growth of ganglion cells, cause atrophy of nerve fibers, and make nerve fibers unable to grow. NSCs, as the basis for the development of mature central nervous system, have the ability for proliferation, differentiation, self-replication, and rapid increase of the number of daughter cells. Like stem cells, the subcellular groups formed by self-replication can self-differentiate and get mature to a certain extent, and the result of differentiation is the formation of various types of cells in the neural tissue. When the brain is damaged, these neural precursor cells can differentiate into neurons and glial cells and play a role. This differentiationability is permanent. NSCs can be reproduced and proliferated under suitable environmental conditions. NSCs can further differentiate and form mature new neurons after replicating to a certain number. However, the adult newborn neurons are unstable and susceptible to various factors. The neurotransmitter, through the conversion of energy, is the medium of transmitting the signal and pulse signal. Furthermore, it can effectively promote NSC proliferation and differentiation of NSC, but the effect of neurotransmitters on NSC changes with the situation, which is different from the normal state of cerebral ischemia.

3. Experiment Set-Up

3.1. Research Object.
Six nests of clean newborn SD rats were divided into three groups. From the 8th day of birth, the model rats were given 60 mmHg of colorectal distention stimulation every afternoon, once a day for seven consecutive days; the lactating rats were raised in the same cage with the mother rats and separated from the mother rats after 21 days; the normal control group was not given colorectal distention stimulation, and other treatments were the same as the model group. The rats were fed with mixed formula for 8 weeks. They can eat and drink freely and change drinking water and feed every day to keep the living environment ventilated and clean. Observe your diet, drinking water, activities, and defecation to ensure their

health and avoid all kinds of harmful stimulation. They were randomly divided into model group, control group, and electroacupuncture group. Rats in the normal control group were treated the same as those in the model group.

3.1.1. Main Instrument Parameters.
The list of main instruments is shown in Table 1.

3.2. Experimental Plan.
The establishment of chronic visceral pain model was done by stimulating Sprague-Dawley rats by colorectal distention once a day when awake, 6–10 days after birth. The CRD method is done by maintaining the pressure for 1 min, deflate, and then retract the balloon. Before expansion stimulation, 0.5 ml physiological water can be used for enema and defecation in advance to reduce the influence of fecal quality on expansion degree. After the modeling, the normal feeding was done for 8 weeks, and the experiment started.

3.2.1. Behavioral Evaluation.
At the age of 8 weeks, the sensitivity of visceral pain in rats was evaluated by the scores of absolute with rawalreflex (AWR) stimulated by CRD. The picture of a rat posing during the CRD experiment is shown in Figure 1.

Before the experiment, the rats fasted for 18 hours. After anesthesia with ether, the self-made wax-coated air bag was inserted into the colorectal part, and the end of the air bag was 1.0 cm deep from the anus, and the catheter was fixed at the place 1.0 cm away from the anus by the adhesive tape. The catheter is connected to a syringe and a sphygmomanometer through a three-way pipe. The rats were placed in a plexiglass observation box. Before and after the experiment, the mice could move freely but could not turn around. The experiment began 30 minutes after the rats woke up and were fully adapted to the environment. Air was injected from low pressure gradient to high pressure gradient, and the pressure in the balloon reached 20 mmHg, 40 mmHg, 60 mmHg, and 80 mmHg, respectively. The expansion time is 20 s, the interval is 4 minutes, and the pressure is repeated three times. The average of the three scores is taken as the result. The single-blind method was used to evaluate the AWR intensity under various pressures. Scoring standard can be given as follows: 0, no obvious behavior change during colon expansion; 1, the body is still or simple head movement during colon expansion; 2, the abdominal muscle began to contract during colon expansion, but the abdominal muscle did not leave the platform; 3, the abdominal muscle obviously contracted to become flat during colon expansion or the lower abdominal wall from the bottom of

FIGURE 1: CRD experiment in rats.

FIGURE 2: Picture of electroacupuncture.

FIGURE 3: AWR score.

the box; 4,+ when colon expansion, the abdominal wall arched or is accompanied by body uplift, pelvis bent, or testicles raised.

3.2.2. Measurement of Extraabdominal Obliquity. When the rats are 6 weeks old, CRD stimulation was used to measure extraabdominal obliquity and evaluate visceral pain sensitivity. After the balloon was implanted as described above, the rats were fixed on the operating table, and the electrode needle was inserted into the lateral oblique muscle 1.5 cm above the inguinal ligament. The experiment began 30 minutes after the rats woke up and were fully adapted to the environment. Air was injected from low to high gradient, and the pressure in the balloon reached 20 mmHg, 40 mmHg, 60 mmHg, and 80 mmHg, respectively. Each expansion lasts for 10 s, and the interval between the two pressures is 4 min. The RM6240BD multichannel physiological signal acquisition and processing system was used to record the discharge activity of rat abdominal external oblique muscle under different CRD pressure (parameter setting: high-frequency filtering 3 kHz, time constant 0.001 s, sampling frequency 40 Hz, sensitivity 500 uv, paper speed 200 ms/div). The experimental environment is quiet, and the relative humidity is within 40–70%, and the room temperature is within 23–27°C.

3.2.3. Electroacupuncture. The acupuncture method was done according to the standard acupuncture method, and the rats were fixed. According to the needs of the experiment, choose Baihui and Dazhui points for electroacupuncture treatment with an electroacupuncture treatment instrument. Baihui is located at the intersection of the middle line of the head and the line connecting the two ear tips, and Dazhui is located under the spinous process of
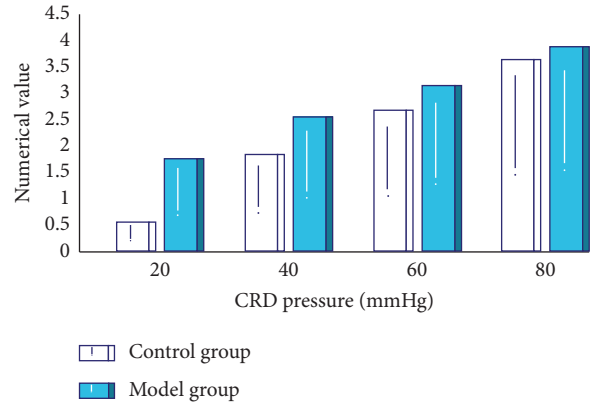
the seventh cervical vertebra. The spinous process of the seven cervical vertebrae is located behind the bowed head under the largest package of the neck. The picture of the electroacupuncture instrument is shown in Figure 2.

3.2.4. Acupuncture Method. Use the 31l inch filiform needle, and after finding the right position, prick the acupoints along the direction of the skin, and the depth of acupuncture is moderate, about 2.5–3.5 mm. The needle handle is connected with the electrode of the electroacupuncture instrument. The strength of the electroacupuncture instrument is adjusted to 3.5 V, and the frequency is 4 Hz. Then close observation is required. The appropriate standard is a slight twitch of the animal's muscles, without struggling and neighing for 10 minutes. The electroacupuncture group can be treated once a day for 14 consecutive days on the day the animal model and after the animal awakes. The density wave was used in the electroacupuncture treatment instrument. CRD was stimulated immediately after electroacupuncture, and the changes of AWR and EMG were observed, and the therapeutic effect of electroacupuncture was evaluated.

3.2.5. Computer Image Analysis. The BX51 digital microscope camera system collects the images of each spinal cord section and magnifies them by 20 times. Under the same multiple, each section randomly selected three fields of vision, and the computer image analysis system was applied to semiquantitative analysis of the grayscale integral value of immune positive cells. The smaller the grayscale integral value, the stronger the tissue staining.

3.2.6. Statistical Analysis. The data were calculated by mean ± standard deviation. The standard of significance test is $a = 0.05$. SPSS11.5 was used for statistical analysis. In order to eliminate the influence of different basic discharges on the experimental results, the basic discharge amplitudes of young rats under the pressure of 20 mmHg, 40 mmHg, 60 mmHg, and 80 mmHgcrd were subtracted from the basic discharge amplitudes of young rats, and the difference represented the discharge amplitudes of different CRDs.

TABLE 2: AWR score.

CRD pressure (mmHg)	Control group	Model group
20	0.6	1.8
40	1.9	2.6
60	2.7	3.2
80	3.7	3.9
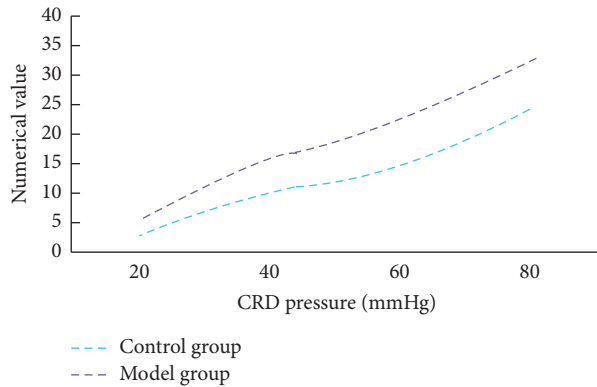
FIGURE 4: Discharge detection of the external oblique muscle.

TABLE 3: Discharge detection of the external oblique muscle.

CRD pressure (mmHg)	Control group	Model group
20	3	6
40	11	17
60	16	25
80	27	36

Repeated ANOVA was used to analyze the AWR score and discharge data.

4. Results

According to the statistical analysis of the data, as shown in Figure 3 and Table 2, with the increasing stimulation pressure, the AWR scores of the model group and the control group increased. In the range of 20–60 mmHg, the AWR score of the model rats was significantly higher than that of the control group ($P < 0.05$), but there was no significant difference in 80 mmHg between the two groups. It suggested that visceral hyperalgesia appeared in the model rats.

According to the statistical analysis of the data, as shown in Figure 4 and Table 3, with the increase of CRD pressure, the discharge amplitude of the external oblique muscle gradually increases. Under 20–80 mmHg CRD pressure, the discharge amplitude of the external oblique muscle of the model rats was significantly higher than that of the control group ($P < 0.05$).

According to statistical analysis of the data, as shown in Figure 5, different CRD pressure stimulation, at the same time, the AWR scores of model rats and normal control rats after electroacupuncture were compared. The results showed different CRD pressures (20–80 mmHg) ($P < 0.05$). The results showed that multiple electroacupuncture had an
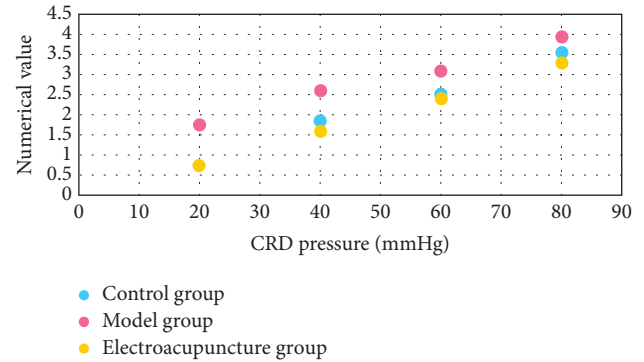
FIGURE 5: Effect of electroacupuncture on visceral pain.
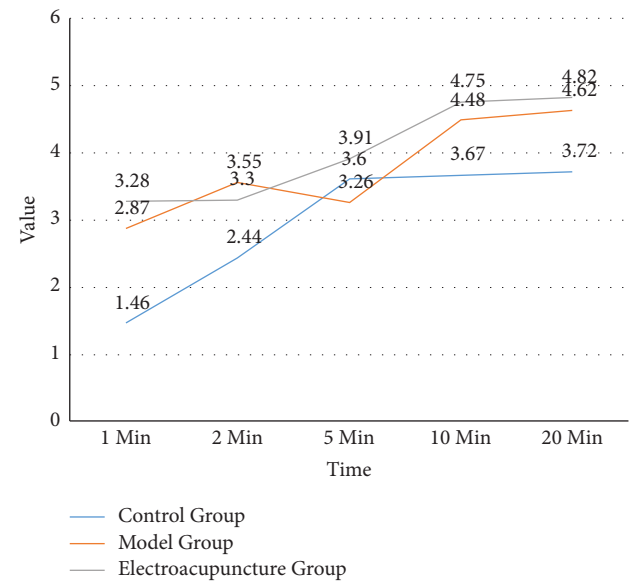
FIGURE 6: The influence of electroacupuncture on EMG.

obvious analgesic effect on the model rats, which could reduce visceral hyperalgesia and restore normal levels.

According to the statistical analysis of the data, as shown in Figure 6, the amplitude of the EMG of model rats was compared with that of the normal control rats after electroacupuncture. The results showed the pressure of 20–80 mmHg ($P < 0.05$). The results showed that multiple electroacupuncture had an obvious analgesic effect in the model rats, which could reduce visceral hyperalgesia and recover to the normal level.

5. Discussion

At present, the pathophysiological process of IBS mainly includes gastrointestinal motility abnormality, visceral sensory disturbance, and gastrointestinal secretion abnormality. In all these pathophysiological processes, 5-hydroxytryptamine (5-HT) was found to be involved. 5-hydroxytryptamine (5-HT) is widely distributed in the peripheral and central nervous system. It is a monoamine with the indoleamine structure. 5-HT can affect its receptor, cause the changes of intestinal perception, movement and

secretion in IBS patients, and produce various symptoms. According to modern medicine, there are not only pain centers but also tissue structures related to analgesia and regulation system in the central nervous system, which can integrate and process all kinds of stimulation information. Acupuncture can play an analgesic role through the integration of some central nerves, body fluids, and pain-regulating systems. The visceral fibers of the descending colon and rectum of the rats project from the pelvic nerve (parasympathetic nerve) to the l6-S2 spinal cord segment, and from the gastric nerve (sympathetic nerve) to the T13-L2 spinal cord segment. The dorsal horn of the spinal cord is the first relay station to transmit harmful information to the central nervous system. After the injury information is transmitted to the dorsal horn, the dorsal horn not only receives and transmits the injury information but also integrates the injury information. The 5-hydroxytryptamine in the spinal cord is derived mainly from the 5-hydroxytryptamine descending projection fiber in the brainstem. These fibers terminate in the dorsal horn of the spinal cord, release 5-hydroxytryptamine, act on receptors in the dorsal horn of the spinal cord, and play a role in pain regulation. There are many serotonin receptor subtypes in the dorsal horn of the spinal cord, which play an important role in the regulation of pain. As an important neurotransmitter, 5-HT2A plays an important role in the transmission and regulation of sensory information, especially nociceptive information at the spinal level. The 5-HT2A receptor plays an important role in the integration of pain information at the spinal level. 5-HT neurons in the brainstem are mainly concentrated in the raphe nucleus. The raphe nucleus group is not only an important mechanism of brain pain control but also a collection of 5-HT neurons, which can send up and down fibers. The ascending 5-HT fibers mainly come from the dorsal raphe nucleus and are distributed in the diencephalon, basal ganglia, and limbic system. Acupuncture information can activate neurons in the dorsal raphe nucleus and inhibit nociceptive neurons in the parafascicular nucleus of the thalamus. The descending 5-HT fibers mainly originate from the nucleus raphe magnus and descend to the spinal cord along the dorsolateral side of the spinal cord. The thalamus is not only an important relay station for transmitting nociceptive information to the cerebral cortex but also a complex analysis and synthesis center. The central central nucleus of the parafascicular nucleus and the plate nucleus is an important center for pain sensation and regulation. The cerebral cortex is the highest center of higher nervous activity, the center of conscious thinking, and the intelligent center of comprehensive reasoning and judgment.

Acupuncture analgesia is not only a simple process of excitation and inhibition but also a complex regulatory process. It not only enhances the analgesic effect but also excessively suppresses the analgesic effect, which is in a dynamic balance. Xu Wei et al observed that peripheral nociceptive stimulation can activate neurons in the sensory cortex, and acupuncture information can also reach the cerebral cortex. The excitatory cortical somatosensory area II can inhibit the nociceptive responses of neurons in the central nucleus, parafascicular nucleus, and central lateral nucleus of the thalamus. However, in the block of somatosensory area II or the change of functional state of somatosensory area and aminobutyric acid II, the inhibitory effect of acupuncture on the pain response was not observed in the atomic nucleus, parafascicular nucleus, and central lateral nucleus of most central intermediate neurons, and it is suggested that the descending activity of the cerebral cortex is involved in the inhibition of pain transmission by acupuncture. Visceral hypersensitivity is related to the imbalance of neurotransmitter release and the increased sensitivity of sensory nerve endings to these media, which may involve many links such as receptors, signal afferent, spinal dorsal horn, and central nerve.; most of the afferent nerve fibers in the gastrointestinal tract are multimodal sensory nerve fibers, which can conduct various stimulation signals. According to histological markers, these fibers can be divided into two groups. One group contains calcitonin gene-related peptide and substance P (SP) and other neuropeptides. CGRP was mainly distributed in the visceral sensory nerve. 50% of sensory neurons containing CGRP also contain neurokinin. By regulating the expression of NKRL in primary afferent neurons of the spinal cord, visceral sensitivity increased. CGRP receptor antagonist can reduce the high sensitivity of colon. CGRP can cause visceral hypersensitivity. CGRP is involved in acupuncture analgesia. The gastrointestinal sensory nerve transmits the stimulation signal to the spinal cord, activates the mechanoreceptor of the primary afferent nerve of the spinal cord, then stimulates the spinal dorsal horn neurons to cause pain, and then transmits the signal to the brainstem. Acupuncture and moxibustion is based on the dialectical principle of "deficiency is tonic and essence is purgative." Through the use of tonic, purgative, flat, purgative, and other techniques, the body's own regulatory response can be achieved. From the perspective of traditional Chinese medicine theory, it is used to regulate meridians and collaterals, qi and blood, and Yin and Yang; from the perspective of modern life science, it plays a role by regulating the function of the neuroendocrine and immune system network. There are many endogenous bioactive substances involved in the change of acupuncture function, which is an active multilevel physiological regulation process.

Nanotechnology is a new subject, which is devoted to the preparation of nanomaterials and the improvement of its practicability in various scientific fields. Metal-doped hybrid materials have been proved to be a drug for many pathogenic microorganisms and various cancers. Among all kinds of biomaterials, bio-based nanometals, including platinum, gold, silver, and their composites with graphene are becoming more and more important. In view of the new significance of metal biomaterials, our research focuses on the preparation of platinum, palladium, zinc oxide, and silver by using natural products as reducing agents and stabilizers of different plants. Second, these metals were loaded onto graphene oxide for clinical biological application. According to traditional Chinese medicine, the large intestine and lungs are located in the abdomen, the stomach meridian of the Foot Yangming, the spleen meridian of the

foot Taiyin, and the meridians and collaterals of the Ren pulse. Therefore, abdominal pain caused by colorectal lesions includes abdominal pain caused by enteroscopy. According to the basic rules of the main treatment of the acupuncture points, the meridian of the acupuncture points such as the large intestine, the stomach meridian of Foot Yangming, the spleen meridian of Foot Taiyin, and the Ren meridian can be selected to relieve pain combined with the commonly used pain treatment methods. In addition to the local protection of intestinal mucosa, NOS can also be used as a reverse transcription information material of visceral sensory components to upload the information of intestinal injury to the relevant parts of the midbrain and spinal cord. According to the theory of somatic visceral convergence, the superficial dorsal horn of the spinal cord will rise along the thalamus tract of the spinal cord, and the gray matter around the central canal will be in the superficial dorsal horn, the medial lemniscus of the posterior line, and reaches the relevant brain area. In the process of nociceptive information transmission, the results of NOS detection are inconsistent, some experimental reports are up-regulated, and some experimental reports are down regulated. The expression of NOS positive neurons in PAG and DH increased after pain stimulation. Under the condition of noxious stimulation, NOS is involved in the transmission and regulation of noxious information at the level of spine and above. This may be due to the fact that in pain in spinal cord, the stimulation causes the release of Glu from the primary afferent terminal, the activation of NMDA receptor, the opening of receptor channels, coupled with the influx of Ca^{2+} and the assistance of calmodulin (CAM), substituting NADPH, catalyzing L-Arg, substituting NO, further activating guanosine cyclase, thus acting on the ion channel of cGMP, the regulation of phosphodiesterase (PDE) by CG liver, and the activation of spinal cord pain system. Many neurotransmitters work through NOS, and the excitatory amino acid NMDA can induce NO synthesis by activating NOS. By regulating the function of neuroendocrine system, acupuncture can improve the body's defense and disease resistance; on the other hand, it is also one of the ways that acupuncture regulates the function of other organs. Generally speaking, acupuncture regulates the neuroendocrine system mainly through the hypothalamic pituitary system and also through the feedback of various neuroendocrine glands. All kinds of pressure and pain will cause over activation of multiple biological cascade systems, which will damage the body. E and NE are the main objective indexes reflecting the changes of neuroendocrine. The change of its content is one of the important indexes reflecting the stress response caused by harmful stimulation.

6. Conclusion

Electroacupuncture can significantly reduce the pain sensitivity of the model rats; electroacupuncture combined with the nerve growth factor can activate the proliferation, migration, differentiation, and integration of endogenous NSC and promote the regeneration and neurological function recovery after nerve injury. Acupuncture can reduce the

visceral pain of colorectal distention by blocking the transmission of pain signals. Acupuncture can cause the changes of NOS, e, NE, and EK. The complex relationship between these neurotransmitters suggests that acupuncture may have a greater application value in improving and alleviating visceral pain due to colorectal swelling. The swelling visceral pain of colorectal in rats was relieved. The blood vessels and nerves in the body are accompanied by similar growth patterns. The activity and nutrition of blood vessels are regulated by nerves and their nutritional factors. NGF is a multifunctional peptide molecule. The amplitude of electromyogram of model rats and normal control rats after electroacupuncture were compared. The results showed the 2080 mmHg group is statistically significant ($P < 0.05$). This shows that NGF promotes the proliferation, migration, and differentiation of endothelial tissue cells while reducing visceral hyperalgesia [26–34].

Data Availability

The data will be available upon the request of the corresponding author.

Conflicts of Interest

There are no conflicts of interest related to this research work and article.

Authors' Contributions

All the authors have collected data and have conducted the study. The drafting is also done by all the authors sectionwise. All have equally contributed to the study.

References

[1] Y. Liu, W. Liu, X. Wang, Z. Wan, Y. Liu, and Y. Leng, "Dexmedetomidine relieves acute inflammatory visceral pain in rats through the erk pathway, toll-like receptor signaling, and Trpv1 channel," *Journal of Molecular Neuroscience*, vol. 66, no. 2, pp. 279–290, 2018.

[2] A. Johnson, C. O. Ligon, R. Latorre, and B. G. V. Meerveld, "Optogenetic activation of central amygdaloid circuitry induces visceral pain in freely moving rats," *Gastroenterology*, vol. 152, no. 5, p. S729, 2017.

[3] M. Kaur, S. R. Sakhare, K. Wanjale, and F. Akter, "Early stroke prediction methods for prevention of strokes," *Behavioural Neurology*, vol. 2022, Article ID 7725597, 9 pages, 2022.

[4] H. Zhu, S. Hu, X. Miao, Y. Xiao, and G. Xu, "Electroacupuncture attenuates visceral pain and reverses upregulation of Trpv1 expression in adult rats with neonatal maternal deprivation," *Chinese Medicine*, vol. 7, no. 1, pp. 1–9, 2016.

[5] M. Kaur, "Elitist multi-objective bacterial foraging evolutionary algorithm for multi-criteria based grid scheduling problem," in *Proceedings of the 2016 Int. Con. On Internet of Things and App. (IOTA)*, pp. 431–436, Pune, India, January 2016.

[6] C. Zhang, Y. Li, X. Wang, Y. Fei, and L. Zhang, "Involvement of neurokinin 1 receptor within the cerebrospinal fluid-contacting nucleus in visceral pain," *Molecular Medicine Reports*, vol. 15, no. 6, pp. 4300–4304, 2017.

[7] C. Dou, "Effect of electroacupuncture and moxibustion on visceral pain and expression of vr 1 and hsp 70 of "tianshu" (st 25) region in colorectal distension-induced visceral hypersensitivity rats," *Acupuncture Research*, vol. 41, no. 4, p. 291, 2016.

[8] N. M. Mule, D. D. Patil, and M. Kaur, "A comprehensive survey on investigation techniques of exhaled breath (EB) for diagnosis of diseases in human body," *Informatics in Medicine Unlocked*, vol. 26, Article ID 100715, 2021.

[9] O. Kilci, T. Demir, M. Günbey et al., "The analgesic effect of diclofenac sodium administered via the epidural route in an experimental visceral pain model," *Nigerian Journal of Clinical Practice*, vol. 19, no. 6, pp. 747–752, 2016.

[10] Y. Kogure, S. Wang, K. I. Tanaka et al., "Elevated h 2 o 2 levels in trinitrobenzene sulfate-induced colitis rats contributes to visceral hyperalgesia through interaction with the transient receptor potential ankyrin 1 cation channel," *Journal of Gastroenterology and Hepatology*, vol. 31, no. 6, pp. 1147–1153, 2016.

[11] Y. Xiao, X. Chen, P. A. Zhang, Q. Xu, H. Zheng, and G. Y. Xu, "Trpv1-Mediated presynaptic transmission in basolateral amygdala contributes to visceral hypersensitivity in adult rats with neonatal maternal deprivation," *Scientific Reports*, vol. 6, no. 1, Article ID 29026, 2016.

[12] Y. Tang, A. Chen, Y. Chen et al., "Zeta inhibitory peptide as a novel therapy to control chronic visceral hypersensitivity in a rat model," *PLoS One*, vol. 11, no. 10, Article ID e0163324, 2016.

[13] M. Kaur and S. S. Kadam, "Discovery of resources using MADM approaches for parallel and distributed computing," *Engineering Science and Technology, an International Journal*, vol. 20, no. 3, pp. 1013–1024, 2017.

[14] L. Liu, Z. A. Cordner, K. L. Tamashiro et al., "The mast cell stabilizer ketotifen attenuates both visceral pain and psychological behavior in a rat model of functional dyspepsia (fd)," *Gastroenterology*, vol. 152, no. 5, p. S928, 2017.

[15] F. Sekiguchi, Y. Kawara, M. Tsubota et al., "Therapeutic potential of RQ-00311651, a novel T-type Ca2+ channel blocker, in distinct rodent models for neuropathic and visceral pain," *Pain*, vol. 157, no. 8, pp. 1655–1665, 2016.

[16] A. Chen, Y. Chen, Y. Tang et al., "Hippocampal ampars involve the central sensitization of rats with irritable bowel syndrome," *Brain and Behavior*, vol. 7, no. 3, Article ID e00650, 2017.

[17] D. K. Prusator and B. Greenwood-Van Meerveld, "Amygdala-mediated mechanisms regulate visceral hypersensitivity in adult females following early life stress: importance of the glucocorticoid receptor and corticotropin-releasing factor," *Pain*, vol. 158, no. 2, pp. 296–305, 2016.

[18] S. P. Lee, O. Y. Lee, K. N. Lee et al., "Effect of da-9701, a novel prokinetic agent, on post-operative ileus in rats," *Journal of Neurogastroenterology and Motility*, vol. 23, no. 1, pp. 109–116, 2017.

[19] K. Nakaya, Y. Nagura, R. Hasegawa, H. Ito, and S. Fukudo, "Dai-kenchu-to, a herbal medicine, attenuates colorectal distension-induced visceromotor responses in rats," *Journal of Neurogastroenterology and Motility*, vol. 22, no. 4, pp. 686–693, 2016.

[20] L. Crispoltoni, A. M. Stabile, A. Pistilli et al., "Changes in plasma β-ngf and its receptors expression on peripheral blood monocytes during alzheimer's disease progression," *Journal of Alzheimer's Disease*, vol. 55, no. 3, 1017 pages, 2016.

[21] A. Ahluwalia, M. K. Jones, N. Hoa, E. Zhu, T. Brzozowski, and A. S. Tarnawski, "Reduced ngf in gastric endothelial cells is one of the main causes of impaired angiogenesis in aging gastric mucosa," *Cellular and Molecular Gastroenterology and Hepatology*, vol. 6, no. 2, pp. 199–213, 2018.

[22] M. Douard, P. Robillard, J. Deweirdt et al., "Connexin-43 expression is increased by the nerve growth factor (ngf) and contributes to pulmonary arterial altered reactivity in pulmonary hypertension," *Archives of Cardiovascular Diseases Supplements*, vol. 10, no. 2, p. 248, 2018.

[23] A. Mahdee, J. Eastham, J. M. Whitworth, and J. I. Gillespie, "Evidence for changing nerve growth factor signalling mechanisms during development, maturation and ageing in the rat molar pulp," *International Endodontic Journal*, vol. 52, no. 2, pp. 211–222, 2019.

[24] D. Luo, J. Zhao, Y. Cheng, S. M. Y. Lee, and J. Rong, "N-propargyl caffeamide (paca) ameliorates dopaminergic neuronal loss and motor dysfunctions in mptp mouse model of Parkinson's disease and in mpp + -induced neurons via promoting the conversion of prongf to ngf," *Molecular Neurobiology*, vol. 55, no. 3, pp. 2258–2267, 2018.

[25] H. T. Do, C. Bruelle, D. D. Pham et al., "Nerve growth factor (ngf) and pro-ngf increase low-density lipoprotein (ldl) receptors in neuronal cells partly by different mechanisms: role of ldl in neurite outgrowth," *Journal of Neurochemistry*, vol. 136, no. 2, pp. 306–315, 2016.

Research Article

Application of Challenging Learning Based on Human-Computer Interaction under Machine Vision in Vocational Undergraduate Colleges

Bin Hu[1,2] and Xueqiong Hong [ID] [2,3]

[1]*Nanchang Business College of Jiangxi Agricultural University, Gongqingcheng 332020, Jiangxi, China*
[2]*Anyang University, Anyang-si 14028, Gyeonggi-do, Republic of Korea*
[3]*Gongqing College of Nanchang University, Gongqingcheng 332020, Jiangxi, China*

Correspondence should be addressed to Xueqiong Hong; hongxueqiong@stu.wzu.edu.cn

Science and technology have progressed in recent years, the deepening of talent education, a challenging learning method of human-computer interaction has gradually emerged. Human-Computer Interaction (HCI for short) is the communication and interaction between humans and machines. This essay aims to apply the challenging learning combined with HCI to vocational undergraduate colleges. The GMM (that is Gaussian mixture model algorithm, commonly used in image recognition or speech recognition, etc.) algorithm is proposed in this essay to recognize students' actions. The effect of HCI is achieved by feeding back the recognized actions to the system. This essay selects 200 students from a vocational undergraduate college for challenging learning (that is, a comprehensive teaching method aiming at students' autonomous learning by stimulating students' interest in learning). The challenging learning designed in this essay is divided into 15 weeks, the task chain contains a total of 196 tasks, and the learning time is 138 h. This essay analyzes the application effects of liberal arts, male and female students, and different grades. The results show that the overall average completion rates of learning tasks for freshman, sophomore, and junior students are about 70%, 75%, and 85%, respectively, and the overall average scores for challenging learning are about 70, 78, and 83. The overall completion rate of weekly tasks for boys and girls is about 68% and 70%, and the overall average score is about 70 points and 75 points. The weekly task completion rate of liberal arts students is generally above 75%, and the overall average score is about 70 points. The overall completion rate of science students is less than 75%, and most of the learning scores are higher than 75 points. In addition, the average accuracy of the GMM algorithm for face and gesture recognition is 90% and 87%. The average frequency of students using HCI is about 320 times a day; the average score of students' experience effect of HCI is about 80 points. It may be stated that the HCI demanding learning strategy proposed in this study worked well and has achieved satisfactory learning results in the application of vocational undergraduate colleges.

1. Introduction

With the gradual deepening of the national talent training plan, the educational methods of vocational undergraduate colleges have also begun to innovate. At present, a challenging learning focusing on cultivating students' abilities is emerging. Challenging learning refers to students' active participation and creativity in challenging thematic learning activities inspired by the real world, emphasizing the development of higher-order thinking, the application of skills, and the internalization of concepts. Challenging learning combines a variety of learning concepts and step-by-step learning methods and conducts periodic learning with goals and tasks. The proposal of challenging learning methods will change the traditional learning methods. The traditional learning method is single, and basically one-to-many undifferentiated learning is carried out according to the course content. The purpose of learning overemphasizes

the acquisition of knowledge, ignoring the cultivation of students' learning style, learning attitude, comprehensive ability, and comprehensive quality. This will lead to students lack of independent thinking, self-directed learning habits, and thinking. The learning ability of each student is different, and teachers cannot adjust the curriculum according to the learning effect of different students. HCI is specialized in studying the way of communication between humans and machines. Combining HCI for challenging learning can make students more interesting in the learning process. It can cultivate different students' comprehensive abilities such as cognitive ability, learning ability, and cooperation ability and play a role in promoting the national talent training plan. Therefore, this essay applies the challenging learning combined with HCI technology to vocational undergraduate colleges, which is of great value and significance.

Regarding the application of challenging learning, some scholars have conducted related research. Among them, Jeffries et al. examined challenging learning strategies in visually impaired (VI) adults. He recognized the importance of engineering and policy education, as well as the relationship between visually impaired mobility demands and new urban initiatives to encourage green mobility [1]. Dillette and Sipe suggested a four-year framework for challenging learning in the context of hospitality and tourist management program leader development. He outlined distinct high-impact events and results that correspond to the development of undergraduate leadership [2]. LA are learning assistants that are used to help students understand content more deeply and are particularly effective in active learning instruction. A fundamental pillar of the LA model is the LA Teaching Curriculum, which teaches LA about evidence-based teaching and how students learn. Purtell et al. used situations in which LA might manifest and reflect in the classroom to provide demanding interactive learning for LA attending pedagogy courses [3]. Puente and Kroesen compared typical college-level courses to information retention and transfer inside a 2nd undergraduate challenging test project. The results showed that students who took the design-based learning program had higher median scores on assignments than students who took the traditional course. The findings of this study serve as a suitable example of the use of design-based learning by higher education practitioners to promote retention and transfer [4]. However, these methods failed to stimulate students' interest in learning, and the learning effect was not high. It is necessary to explore high-quality methods.

To address the issues with students' lack of learning interest and low learning efficiency, some scholars have proposed methods using human-computer interaction. Among them: El Said reports thematic analysis of semi-structured interviews with university students in 52 smart learner interfaces [5]. Ciechanowski et al. studied interpersonal contact, concentrating on users' emotional responses to various sorts of devices they engage with. Understanding the user side, according to Ciechanowski et al., might be crucial in building improved automation throughout the next. As a result, it can contribute to the advancement of the field of HCI [6]. Long and Aleven

thought that intelligent tutoring systems (ITS) and human-computer interactive educational games assist learning by doing, and he assessed students' formula-solving abilities [7]. Chakraborty's et al. research investigated the main limitations of vision-based gesture recognition that arise in detection and preprocessing, representation and feature extraction, and recognition and explored current challenges in detail [8]. Although these methods have improved students' interest in learning to a certain extent, the actual effect is still not very satisfactory, so it needs to be further improved.

This essay adopts the GMM algorithm of HCI to study the challenging learning and proposes a challenging learning method based on HCI. This essay uses machine vision technology to obtain students' action images, and then uses the GMM algorithm and Human 3.6 M action database to recognize the images and feeds the recognition content to the system to realize HCI. This essay picks 200 individuals at random from a college for challenging learning tests. The findings indicate that the challenging learning approach proposed in this essay will have different application effects for different grades, arts and sciences, and between male and female students, but all of them can achieve good learning effects. In addition, the GMM algorithm has a high recognition effect, and this essay's HCI system received a positive rating from the students.

2. HCI-Based Challenging Learning Approach

2.1. Challenging Learning. Challenging learning is a type of heuristic learning, which is a step-by-step learning strategy that helps to improve students' thinking and learning ability, etc. [9]. Figure 1 depicts the HCI challenge learning framework developed in this essay, which consists of a challenge learning scaffold, an application layer, and an HCI layer. The challenge learning scaffold consists of recurrent learning sheets, task chains, and project books [10]. The task chain is the deconstruction of the entire project and the superposition of various related tasks, while the cyclic learning alone is the learning of challenging tasks with time regularity. A project book usually refers to a topic of study and a research project, during which students progress through the learning process, culminating in a challenging learning master project assignment. The application layer is to develop the challenging learning scaffold into a learning software or website through a programming language so that students can learn through mobile phones, computers, tablets, and other devices [11]. This article uses java programming to build a learning website so that students can learn anytime and anywhere on the website. The HCI layer is the design of the way of communication between humans and machines. During the learning process, students interact with machines to make learning more fun [12].

The traditional challenging learning is just to build a simple challenging learning scaffold, and teachers use the learning scaffold to challenge students to learn. Although this method can stimulate students' interest in learning to a certain extent, the teaching method is still in the form of one-to-many teachers, which cannot carry out differentiated
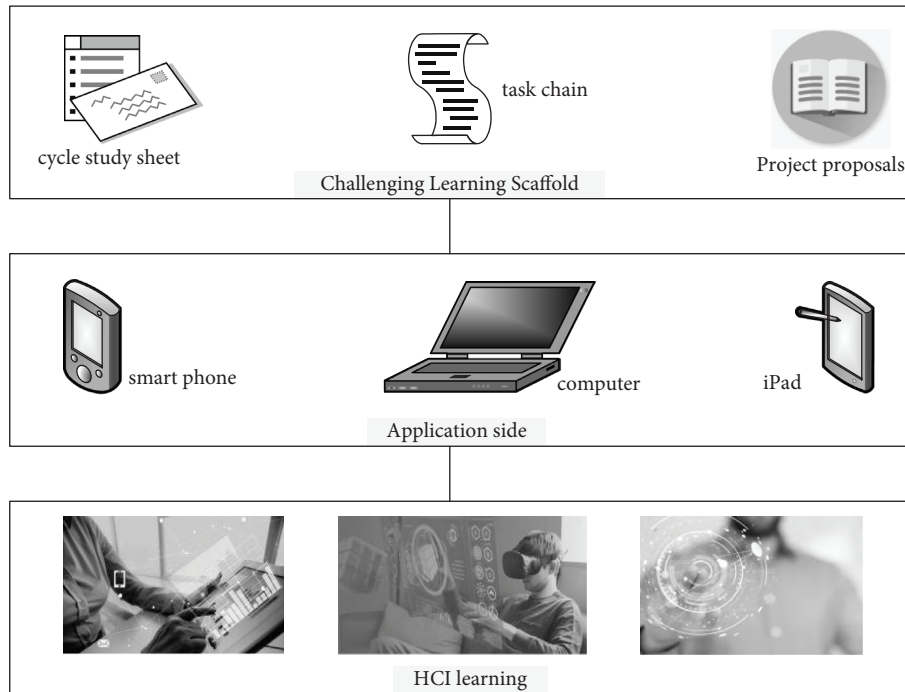
FIGURE 1: HCI-based challenging learning framework.
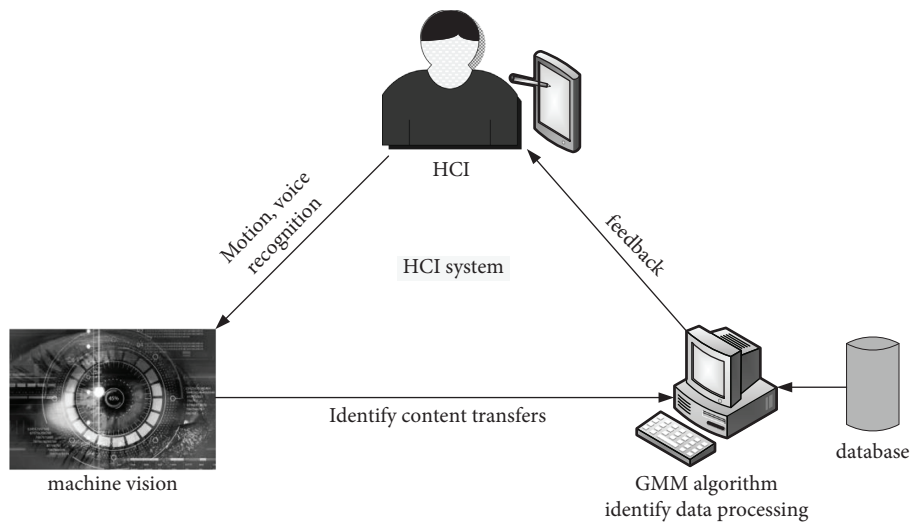
FIGURE 2: HCI system design under machine vision.

teaching, and the students' self-challenging learning efficiency is not high [13, 14]. The HCI challenging learning method designed in this essay can carry out differentiated teaching according to the different learning abilities of students and can further stimulate students' interest in learning. However, the current challenging learning's understanding of "learning" is still relatively narrow and limited, and it is more inclined to the cognitive level of psychology, and students' value, attitude, emotion, and spiritual learning are ignored [15, 16]. The reason for challenging learning is to allow students to use familiar technology and take practical problems related to students as the starting point and foothold in today's era of rapid

development of information technology, which can stimulate students' learning fun and tap their potential [17, 18].

2.2. Design of HCI. HCI is the interaction between human and machine. The design of HCI should be optimized from human touch, vision, and hearing so that users can get a more comfortable experience effect [19]. Figure 2 depicts the HCI system developed in this essay using a robotic system. During the learning process of students, the camera under machine vision collects information such as students' actions, behavior, and speech. The computer receives this information and converts it into data for processing. In this

TABLE 1: Students involved in challenging learning.

Gender	Boys	Girls
Freshman	46	32
Sophomore	49	30
Junior	25	18
Total	120	80

TABLE 2: Challenging learning arrangements.

Period	Number of tasks	Total study duration:h	The average duration of the task:h
Week 1	7	7	1.00
Week 2	7	7	1.00
Week 3	8	7.5	0.94
Week 4	8	7.5	0.94
Week 5	10	8	0.80
Week 6	10	8	0.80
Week 7	12	9	0.75
Week 8	12	9	0.75
Week 9	15	10	0.67
Week 10	15	10	0.67
Week 11	15	10	0.67
Week 12	19	11	0.58
Week 13	19	11	0.58
Week 14	19	11	0.58
Week 15	20	12	0.60
Total	196	138	0.70
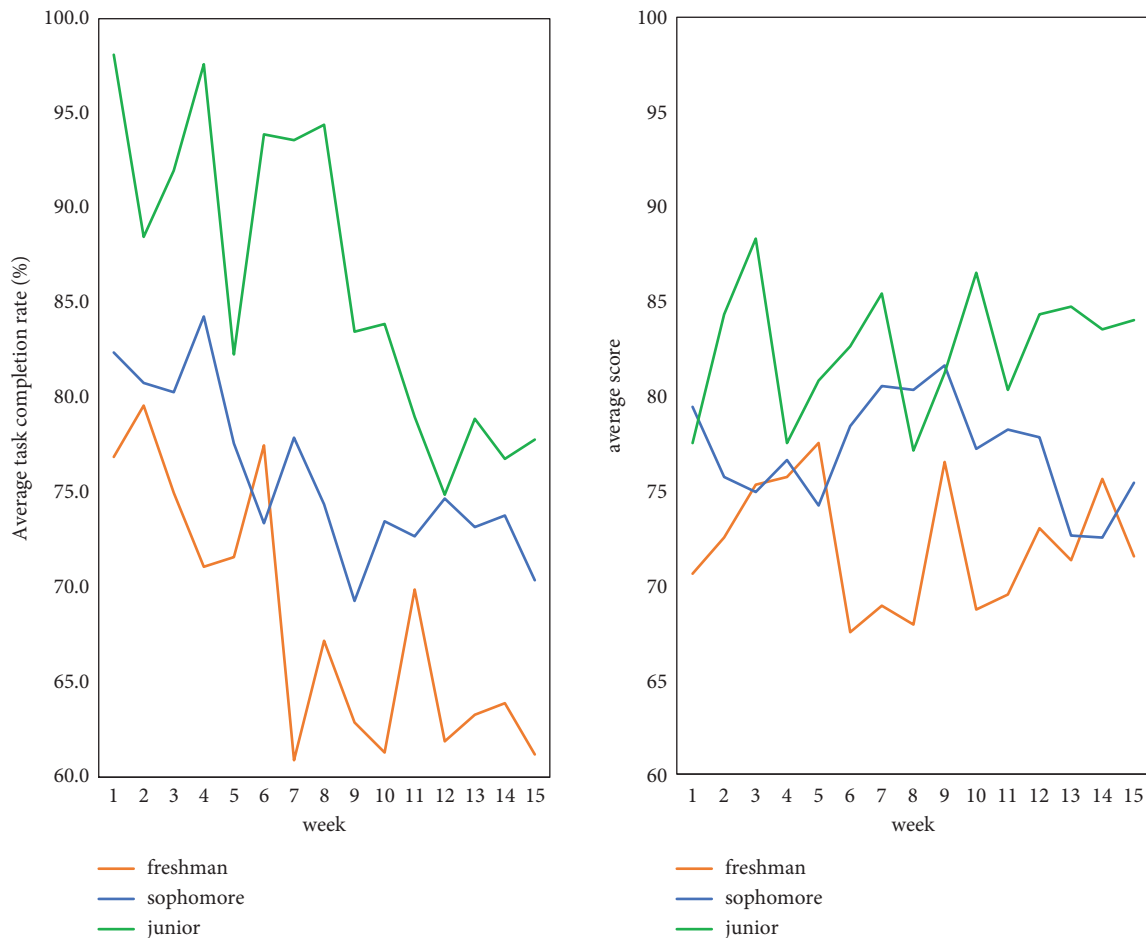
FIGURE 3: Application effect of students in different grades.
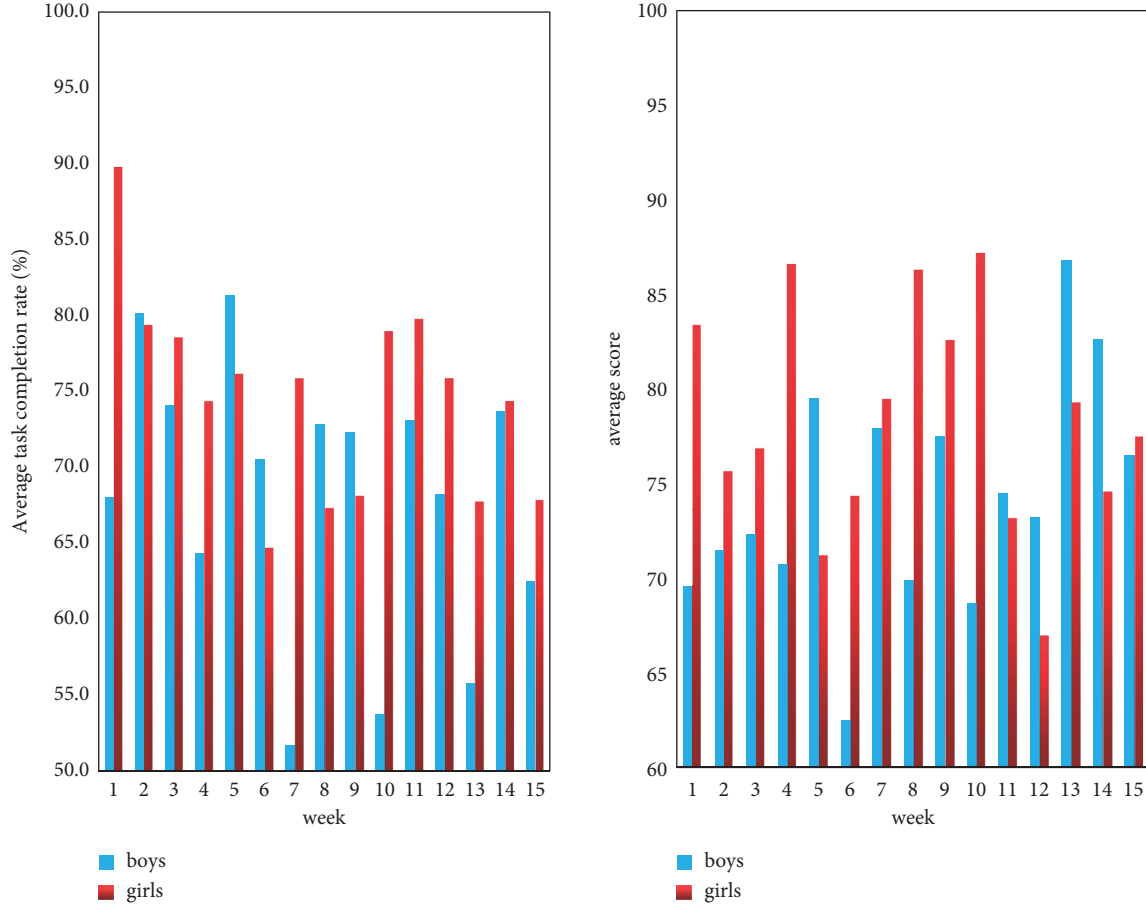
FIGURE 4: The application effect of boys and girls.

essay, the GMM method is utilized to identify and process data. The GMM algorithm uses a probability algorithm, so a data point can belong to multiple clusters, which means that the GMM algorithm can support mixed attributes. And, then feed back to the application according to the result of the identification, so that the machine can respond accordingly, thus realizing the process of HCI.

The design of HCI should follow the principles of simplicity, beauty, and convenience, bringing fun and comfort to people. The goal of this essay's HCI system is to use machine vision to recognize human behaviors. Then, the meaning of the behavior is analyzed according to the human behavior database, and the analysis result is returned to the program, and the application side makes matching interaction information [20]. This method can identify students' behaviors in real time and give more accurate feedback information in the challenging learning process.

2.3. GMM Algorithm.
The GMM algorithm is an algorithm for unsupervised learning using a Gaussian model (a probability model based on normal distribution, often used in data analysis and image processing). In this essay, the GMM algorithm is used to identify students' action behaviors in challenging learning, so as to achieve the effect of HCI. The algorithm is identified as follows.

Assuming that the image is I, its pixels are represented by a Gaussian model, and its single initial model is

$$
\begin{cases}
\mu(c, g, 0) = I(c, g, 0) \\
\sigma^2(c, g, 0) = \mathrm{std}_{\mathrm{init}^2}, \\
\sigma(c, g, 0) = \mathrm{std}_{\mathrm{init}}.
\end{cases}
\tag{1}
$$

Among them, μ is the mean, the deviation is denoted by ρ, and ρ^2 is the variance.

The output image is represented by Iop, then the detection result is

$$
Iop(x, y, t) =
\begin{cases}
0, & |I(c, g, t) - \mu(c, g, t)| < \lambda \times \sigma(c, g, t - 1), \\
1, & \text{otherwise}.
\end{cases}
\tag{2}
$$

Among them, λ is the adjustment parameter.

With the change of time, the input and output images need to be updated, and the update method is as follows:

$$
\begin{cases}
\mu(c, g, t) = (1 - a) \times \mu(c, g, t - 1) + a \times \mu(c, g, t), \\
\sigma^2(c, g, t) = (1 - a) \times \sigma^2(c, g, t - 1) \\
\quad + a \times |I(c, g, t) - \mu(c, g, t)|^2, \\
\sigma(c, g, t) = \sqrt{\sigma^2(c, g, t)}.
\end{cases}
\tag{3}
$$
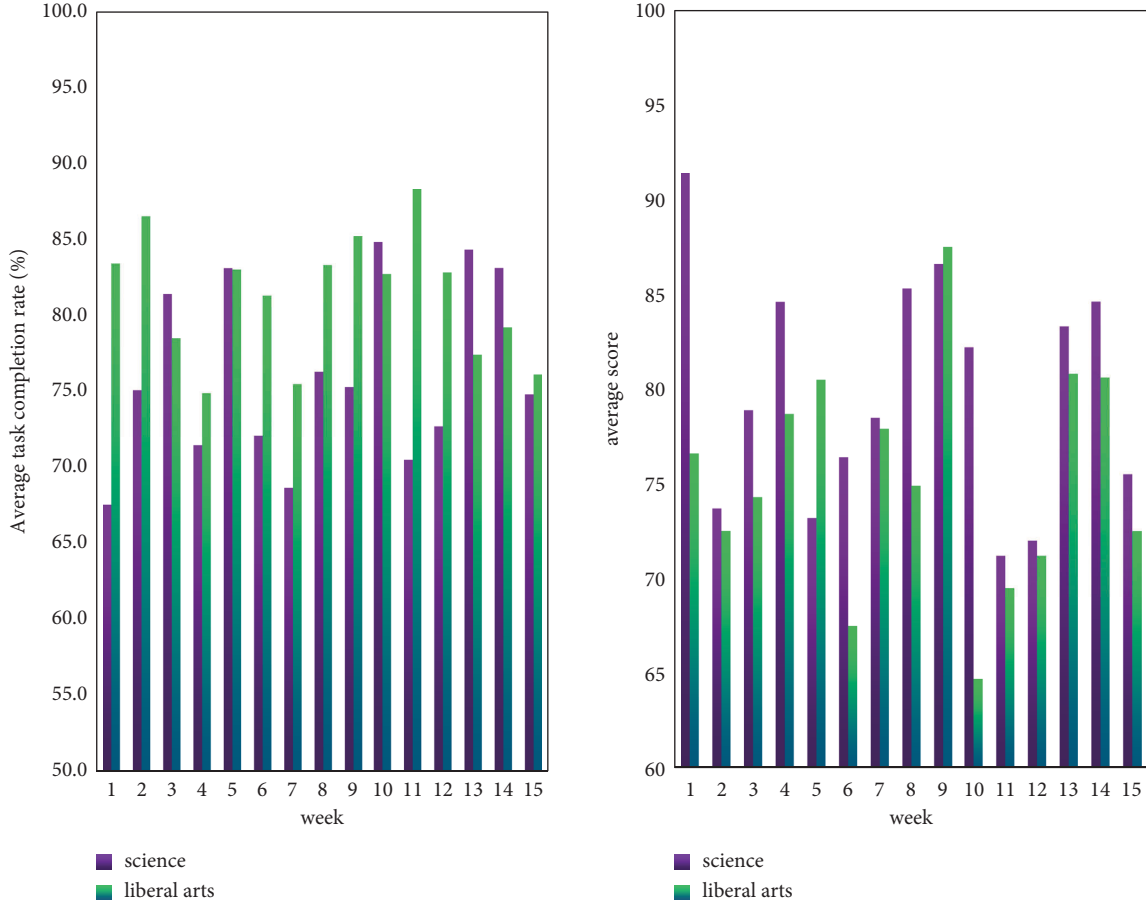
FIGURE 5: Application effect of arts and sciences.

Among them, a is the update rate.

The recognition accuracy and speed of a single model are insufficient and often cannot meet the inspection requirements. Therefore, multiple Gaussian models are often used for calculation. For the mixed Gaussian model,

$$P(p) = \left[\omega_i(c,g,t), \mu_i(c,g,t), \sigma_i(c,g,t)^2\right]. \quad (4)$$

Among them, $\omega_i(c,g,t)$ is the weight of either model, and

$$\sum_{i=1}^{K} \omega_i(x,y,t) = 1. \quad (5)$$

Among them: K is the number of models.

If d is used to represent the weight increment, it is

$$d\omega = a \times (1 - \omega_i(c,g,t-1)). \quad (6)$$

Then, the weighted weight is

$$\omega_i(c,g,t) = \omega_i(c,g,t-1) + d\omega. \quad (7)$$

Normalize the weights (The purpose of normalization is to facilitate comparison calculations and make the calculated values simpler):

$$\omega_i(c,g,t) = \frac{\omega_i(c,g,t)}{\sum_{i=1}^{K} \omega_i(c,g,t)}. \quad (8)$$

Use the ratio of weight to standard deviation as the basis for model ranking, namely,

$$\text{sort}_{\text{key}} = \frac{\omega_i(c,g,t)}{\sigma_i(c,g,t)}. \quad (9)$$

From this, the image recognition result of the mixture Gaussian model can be obtained.

2.4. Summary of Data Sources. The development language used in this article is java, because this programming language is relatively simple, widely used, and has strong compatibility. The development software is OpenCV, because the software contains java language interface, and the software's visual processing algorithm is powerful and the processing speed is very fast. The visual recognition database is Human 3.6 M, which contains more than 3 million three-dimensional human poses, which can meet the recognition requirements of this essay. In addition, this essay randomly selects 200 students from a vocational undergraduate college for challenging learning. The selected students include freshmen, sophomores, juniors, and students from different majors.Table 1 depicts the specific number [21].

In addition, the challenging learning designed in this essay is divided into 15 weeks, the task chain contains a total of 196 tasks, and the learning time is 138 h. As shown in Table 2, in order to allow students to gradually adapt to this
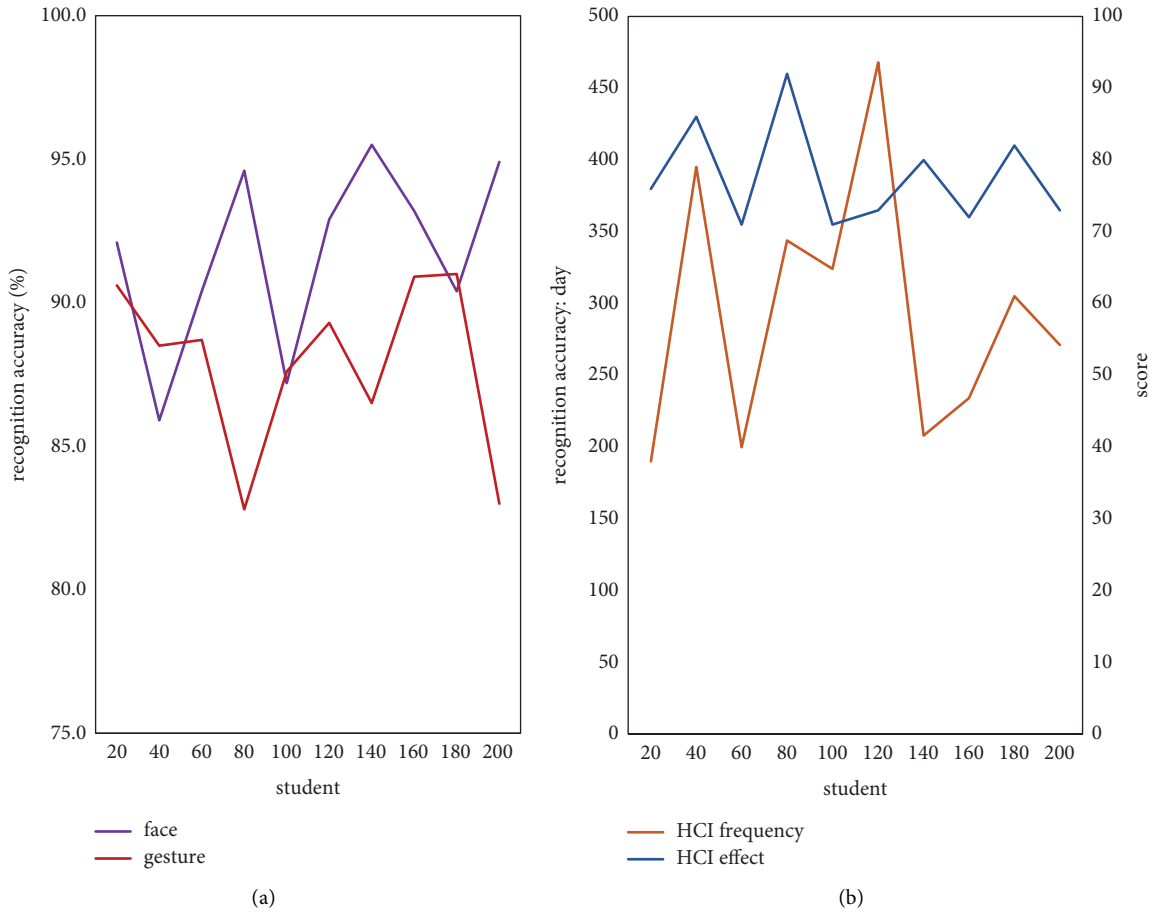
FIGURE 6: Performance analysis of the system application. (a)The recognition effect of GMM algorithm. (b) The use effect of HCI.

3. Results and Discussion

3.1. Application Deconstruction of Students of Different Grades. Through these 15 weeks of challenging learning, this essay makes statistics based on the students' task completion rate and teachers' scores on students' completion effects, as shown in Figure 3.

From Figure 3, it can be seen that the completion rate of freshman students is between 60% and 80%, the overall average completion rate is less than 70%, the learning score is 68–78 points, and the overall average score is about 70 points. Sophomores have a 70% to 85% completion rate, with an overall average completion rate of about 75%, and a learning score between 72 and 82, with an overall average score of about 78. Completion rates for juniors ranged from 75% to 98%, with an overall average completion rate of over 85%, with ratings ranging from 78 to 88, with an overall average rating of 83. This shows that challenging learning will be more effective for students in higher grades. This may

learning rhythm, only 7 tasks and 7 study hours are arranged in the first and second weeks, and the average study time for one task is one hour [22, 23]. Then, the sum of tasks gradually increases, the learning time increases, and the average learning time for each task decreases.

be because students in the upper grades are more pressured to study and are more willing to take the time to study.

3.2. Application Deconstruction of Male and Female Classmates. This essay does data analysis upon this proposal of male and female students to reflect the application influence of various things, as shown in Figure 4.

As can be seen from Figure 4, the weekly task completion of girls is basically higher than that of boys. The highest completion rate is 90%, the lowest is about 64%, and the overall completion rate is above 70%. The highest learning score is 87 points, the lowest is 67 points, and the overall average score is about 75 points. The weekly completion rate of boys is between 52% and 82%, and the overall completion rate does not exceed 68%. The reason for this may be that boys are more playful and spend less time on learning. Boys scored between 63 and 86 on learning, with an overall average of around 70.

3.3. Applied Deconstruction of Arts and Sciences. In addition, this essay divides students into liberal arts majors and science majors according to different majors and conducts data analysis on these two categories, as shown in Figure 5.

It can be seen from Figure 5 that most of the weekly tasks in the liberal arts are better than those in the sciences, and the completion rate is generally above 75%. However, the score of learning is not as high as that of science, with scores ranging from 65 to 88, and the overall average score is about 70 points. The weekly completion rate of science students was between 65% and 84%, and the overall completion rate was lower than 75%; but most of the learning scores were higher than 75 points. It can be seen that, compared with science students, challenging learning can attract more interest in liberal arts students, but the learning effect is actually not as good as that of science students. The reason may be that liberal arts students are more willing to spend time studying, but their thinking ability is not as flexible as that of science students, so the learning effect is not as good as that of science students.

3.4. System Performance Deconstruction for Applications. Finally, this essay collects the recognition results of students' faces and gestures according to the GMM algorithm during the challenging learning process and counts the frequency of students' HCI and the response effect of HCI. Among them, the response effect of HCI is that students score according to their own experience and feelings, and the full score is calculated as 100 points. The results are shown in Figure 6.

It can be seen from Figure 6 that the recognition accuracy of the GMM algorithm for faces is between 86% and 96%, and the average accuracy rate is higher than 90%. The recognition accuracy of gestures will be slightly lower, between 83% and 92%, and the average accuracy is about 87%. The recognition effect is not bad, but there is still room for improvement. In addition, students use HCI at a maximum of 480 times a day and a minimum of 230 times, with an average frequency of about 320 times. The students' scores for HCI experience are between 70 and 90 points, and the average score is about 80 points, which shows that the HCI method designed in this essay can be loved by most students and can achieve good practical application results.

4. Conclusion

As a new way of learning, challenging learning can stimulate students' interest in learning and improve learning efficiency, so it should be applied to vocational undergraduate colleges as soon as possible. Over the years, with said advancement of device vision technology, people have found HCI to be an efficient and interesting way to interact. Therefore, this essay first studies the specific content of challenging learning and HCI and finds that the key to challenging learning is to build a challenging learning scaffold, and the design of the HCI system requires the use of multiple recognition techniques. Then, this essay proposes a GMM recognition algorithm, applies it to the recognition of HCI systems and constructs a challenging learning system based on HCI. Secondly, the learning system was tested. From the grades, genders, and completion of challenging learning in the arts and sciences, it was known that the students' completion of learning tasks and scores were good.

From the system performance test, it is concluded that the recognition accuracy of the GMM algorithm has become more sophisticated, and the students' perception of the use of HCI has become more complex. Therefore, the challenging learning system designed in this essay has obtained decent results in the application of vocational undergraduate colleges. However, the authors acknowledge that their capabilities are limited, and the research has much room for improvement. The author's research on HCI is insufficient, the experiment is not perfect, and the author still needs to conduct more in-depth research and exploration and strive to improve in future work.

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] J. M. Jeffries, R. Gilroy, and T. Townshend, "Challenging the visual: learning from the mobility narratives of visually impaired persons," *Journal of Urban Design*, vol. 25, no. 2, pp. 254–274, 2018.

[2] A. Dillette and L. Sipe, "A systematic framework of experiential learning: challenging educators to make college more than an aggregation of credits," *Creative Education*, vol. 09, no. 09, pp. 1426–1443, 2018.

[3] A. Purtell, R. M. Talbot, and M. E. Moore, "Barriers to learning assistant engagement: an investigation into student encounters learning assistants find challenging and developing training to navigate those challenges," *Journal of College Science Teaching*, vol. 49, no. 6, pp. 23–29, 2020.

[4] S. M. G. Puente and G. M. W. Kroesen, "Facilitating retention and transfer of physics concepts with challenging assignments in design-based learning projects," *Open Journal of Social Sciences*, vol. 08, no. 12, pp. 366–387, 2020.

[5] G. R. E. Said, "The intention to use mobile student portal: a mobile human computer interaction study in a university context in Egypt," *Mobile Information Systems*, no. 2, pp. 1–8, 2018.

[6] L. Ciechanowski, A. Przegalinska, M. Magnuski, and P. Gloor, "In the shades of the uncanny valley: an experimental study of human-chatbot interaction," *Future Generation Computer Systems*, MAR, vol. 92, pp. 539–548, 2018.

[7] Y. Long and V. Aleven, "Educational game and intelligent tutoring system: a classroom study and comparative design

analysis," *ACM Transactions on Computer-Human Interaction*, vol. 24, no. 3, p. 1, 2017.

[8] B. K. Chakraborty, D. Sarma, M. K. Bhuyan, and K. F. Macdorman, "Review of constraints on vision-based gesture recognition for human-computer interaction," *IET Computer Vision*, vol. 12, no. 1, pp. 3–15, 2018.

[9] S. Ghosal, D. Blystone, A. K. Singh, B. Ganapathysubramanian, A. Singh, and S. Sarkar, "An explainable deep machine vision framework for plant stress phenotyping," *Proceedings of the National Academy of Sciences*, vol. 115, no. 18, pp. 4613–4618, 2018.

[10] L. Y. Chang, S. P. He, Q. Liu, J. l. Xiang, and D. F. Huang, "Quantifying muskmelon fruit attributes with A-TEP-based model and machine vision measurement," *Journal of Integrative Agriculture*, vol. 17, no. 6, pp. 1369–1379, 2018.

[11] J. M. Lee, I. H. Jung, and K. Hwang, "Classification of beef by using artificial intelligence," *Journal of Logistics, Informatics and Service Science*, vol. 9, no. No. 1, pp. 1–10, 2022.

[12] Y. Son and Y. S. Lee, "A smart contract weakness and security hole analyzer using virtual machine based dynamic monitor," *Journal of Logistics, Informatics and Service Science*, vol. 9, no. No. 1, pp. 36–52, 2022.

[13] S. Amatya, M. Karkee, Q. Zhang, and M. D. Whiting, "Automated detection of branch shaking locations for robotic cherry harvesting using machine vision," *Robotics*, vol. 6, no. 4, p. 31, 2017.

[14] F. You, Y. H. Li, L. Huang, K. Chen, R. H. Zhang, and J. M. Xu, "Monitoring drivers' sleepy status at night based on machine vision," *Multimedia Tools and Applications*, vol. 76, no. 13, pp. 14869–14886, 2017.

[15] Q. Wang, B. Chen, and D. Zhu, "Machine vision-based selection machine of corn seed used for directional seeding," *Nongye Jixie Xuebao/Transactions of the Chinese Society of Agricultural Machinery*, vol. 48, no. 2, pp. 27–37, 2017.

[16] S. W. Ju and Y. S. Park, "Public design method based on smart service system technology: centered on the cases of bus stops in Korea and China," *Journal of Logistics, Informatics and Service Science*, vol. 9, no. No. 1, pp. 177–194, 2022.

[17] D. M. Tsai and Y. C. Hsieh, "Machine vision-based positioning and inspection using expectation-maximization technique," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 11, pp. 2858–2868, 2017.

[18] Y. Min, B. Xiao, J. DaNg, B. Yue, and T. Cheng, "Real time detection system for rail surface defects based on machine vision," *EURASIP Journal on Image and Video Processing*, no. 1, pp. 3–11, 2018.

[19] M. Abdollahpour, M. R. Golzarian, A. Rohani, and H. Abootorabi Zarchi, "Development of a machine vision dual-axis solar tracking system," *Solar Energy*, vol. 169, no. jul, pp. 136–143, 2018.

[20] P. Kunakornvong and P. Sooraksa, "A practical low-cost machine vision sensor system for defect classification on air bearing surfaces," *Sensors and Materials*, vol. 29, no. 6, pp. 629–644, 2017.

[21] A. A. Robie, K. M. Seagraves, S. E. R. Egnor, and K. Branson, "Machine vision methods for analyzing social interactions," *Journal of Experimental Biology*, vol. 220, no. 1, pp. 25–34, 2017.

[22] T. Dawood, Z. Zhu, and T. Zayed, "Machine vision-based model for spalling detection and quantification in subway networks," *Automation in Construction*, vol. 81, no. sep, pp. 149–160, 2017.

[23] C. Kavitha and S. D. Ashok, "A new approach to spindle radial error evaluation using a machine vision system," *Metrology and Measurement Systems*, vol. 24, no. 1, pp. 201–219, 2017.

Retraction

Retracted: Visualization Method of Key Knowledge Points of Nursing Teaching Management System Based on SOM Algorithm and Biomedical Diagnosis

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 L. Fang, "Visualization Method of Key Knowledge Points of Nursing Teaching Management System Based on SOM Algorithm and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7057437, 9 pages, 2022.

Hindawi

Research Article

Visualization Method of Key Knowledge Points of Nursing Teaching Management System Based on SOM Algorithm and Biomedical Diagnosis

Die Hu [1] **and Le Fang**[2]

[1]*Department of Nursing, Zhengzhou Health Vocational College, Zhengzhou 450122, China*
[2]*Ophthalmology Department, People´s Hospital of Zhengzhou, Zhengzhou 450122, China*

Correspondence should be addressed to Die Hu; 194630318@smail.cczu.edu.cn

The traditional nursing teaching knowledge point recommendation algorithm based on collaborative filtering is difficult to deal with the problem of data sparsity, while the traditional recommendation algorithm based on matrix decomposition has poor scalability in dealing with high-dimensional data, and their recommendation results are only determined according to the prediction score, resulting in low recommendation accuracy. In view of this, a nursing teaching knowledge point recommendation method based on a SOM neural network and ranking factor decomposition machine is proposed. Firstly, the SOM neural network is used to cluster users based on users' academic background information, then the partial order relationship of nursing teaching knowledge points is constructed by using users' explicit and implicit web access behavior, and finally, the factor decomposition machine is used as the ranking function to classify users' academic background web access behavior, borrowing nursing teaching introduction text, and other characteristic information were modeled, and the peer-to-peer ranking learning algorithm was used to accurately recommend nursing teaching knowledge points. Experimental results show that the proposed method can effectively alleviate the problem of data sparsity and improve the accuracy and efficiency of recommendations.

1. Introduction

With the continuous advancement of the digital construction of nursing teaching knowledge points, the number of electronic nursing teaching knowledge points has increased sharply, resulting in problems such as information overload and cognitive loss when users search for nursing teaching knowledge points [1]. Therefore, how to provide users with nursing teaching knowledge point recommendation services according to the users' preferences for nursing teaching knowledge points has become an important problem to be solved to improve the personalized service quality of nursing teaching knowledge points. Most of the existing personalized nursing teaching knowledge point recommendations are realized by the traditional recommendation method based on user collaborative filtering. The basic principle is to find the nearest neighbor users similar to the target users by

calculating the similarity between users, then predict the score value of the target users on the nursing teaching knowledge points according to the historical score data of the similar users on the nursing teaching knowledge points, and recommend the nursing teaching knowledge points based on the score value. Because in the case of sparse data, a user-based collaborative filtering algorithm [2]. Therefore, most scholars are committed to improving the above algorithms. For example, song Chuping integrates reader characteristics and nursing teaching characteristics into user similarity calculation to improve the accuracy of recommendation [3]. However, with an increase in the number of users, the amount of user similarity calculation will increase, resulting in a reduction in recommendation efficiency. Therefore, the SOM algorithm is used for clustering. By calculating the similarity between the target user and each clustering center, the cluster is found, and the nearest

neighbor user set is constructed, so as to reduce the amount of user similarity calculation. However, because the traditional algorithm is affected by the initialization K value and the clustering time is long, the accuracy of clustering results is not high [4].

2. Visualization of Key Knowledge Points of Nursing Teaching Management System

2.1. SOM Structure of Key Knowledge Points of Nursing Teaching Management System. A nursing knowledge map is a specific knowledge base for the nursing field, which includes a series of entities in the nursing field and their associations. There is a precedent for the construction of a nursing knowledge map [5]. Using the technologies of text extraction, relational data conversion, and data fusion, this paper explores the automatic construction method and standardized process of a TCM knowledge map in order to realize the template-based TCM knowledge Q&A and the auxiliary prescription based on the knowledge map reasoning [6]. The SOM network structure is shown in Figure 1.

A SOM network is a tissue feature mapping network. Its basic principle is that for each input vector, a neuron in the output layer has the closest value to the input vector and wins by receiving the maximum stimulation. Some neurons around the winning neuron are also greatly stimulated due to lateral action. At this time, the network performs a learning operation, and the winning neuron and its surrounding neurons modify their own weight vector to move to the input vector [7]. Each neuron is moved to the whole input space as more vectors are submitted, and it is close to the input vector of its nearest vector value and arranged in that layer to obtain a classification. According to experience, when all input and output values are between 0 and 1, the calculation effect of the SOM neural network is the best [8]. Assign the weights of each neuron in the network to the random number in the [0,1] interval as the initial value w_{ij}, set a large neighborhood radius n, and set the number of neurons that are learning. Randomly select a training mode $x(t) = (x_1(t), x_2(t), ..., x_n(t))$ to provide to the input layer of the network. Select the neuron matching the input vector as the winning neuron c. If the Euclidean distance is adopted, C is

$$\|x(t) - w_c(t)\| = \min\left\{\|x(t) - w_j(t)\|\right\}. \tag{1}$$

Update the weights of neurons in the neighborhood to make them move in the direction of super input vector. Assuming that an n-dimensional input eigenvector can be expressed as $x = (x_1, x_2, ..., x_n) \in R_n$, y_i as the target prediction value corresponding to the input eigenvector, FM can use the decomposition interaction parameters to model all nested interactions of n input variables of x in d-dimension [9]. When $d = 2$, the factorization machine model can be expressed as follows:

$$\hat{y}(x) := w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} w_{i,j} x_i x_j. \tag{2}$$
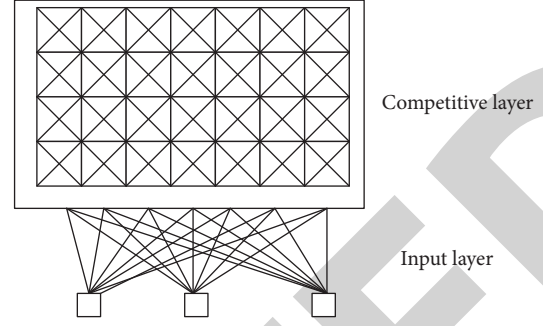
Figure 1: The SOM network structure.

Where w_0 is the total deviation; w_i is the unary interaction parameter of the input variable x_i; w_{ij} is the decomposition parameter between V_I and V_J, which is defined as follows:

$$w_{i,j} = \langle v_{in} v_j \rangle = \sum_{f=1}^{k} v_{i,f} \cdot v_{j,f}. \tag{3}$$

Where k is a super parameter that defines the decomposition dimension. Suppose there is an input space $X \in R_N$, where n is the number of features. At the same time, there is an output space (i.e. scoring space) in which the tag $Y = \{r_1, r_2, ..., r_q\}$ represents the user's preference order for items, and the fixed order they maintain is $r_q, r_{q-1}, ..., r_1$ where v represents the preference relationship. In order to determine the order relationship between items, we need to select a set of sorting functions $f \in F$ so that each candidate function f in f can determine the following partial order relationship, namely:

$$x^{(i)} > x^{(j)} \Longleftrightarrow f\left(x^{(i)}\right) > f\left(x^{(j)}\right). \tag{4}$$

Suppose, in $x \times$ there is a set of sorting reals $S = \{(x(i), y(i))\}$, $T_i = 1$ in y space, where $y(i)$ is the preference sorting, and t is the number of instances. The sorting task is to find an optimal function $f * \in f$ to minimize the loss function of a given sorting instance. Here, F_m function is selected as the sorting function, which is as follows:

$$f_\theta(x) := w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{ji+1}^{n} x_i x_j \sum_{f=1}^{k} v_{I,f} v_{j,f}. \tag{5}$$

Convert any instance pair and their sequential relationship into a new instance, and give the instance a new label. Assuming that p and q, respectively, represent an instance in an instance pair, and y_p and y_q represent their sorting, there are

$$(p, q), z = \begin{cases} +1 & y_p > y_q, \\ -1 & y_\phi > y_j. \end{cases} \tag{6}$$

According to the above method, a new training set s' = {P(t), q(t), and Z(t)}, $L_i = 1$, can be created from a given training set s, where l is the number of newly constructed instances. Thus, the hinge loss function of the t-th instance pair in the training set s' is

$$l_i\left(f; p^{(1)}, q^{(0)}, z\right) = \left[1 - z \times \left(f_\Theta\left(p^{(\omega)}\right) - f_\Theta\left(q^{(\theta)}\right)\right)\right]_+. \quad (7)$$

Where the subscript "+" represents the positive part; f(P(t)) - f(q(t)) can be calculated by FM function within the linear time complexity $O(k\bullet n)$. Define a global loss function on the whole training set s':

$$\min_\Theta L(\Theta) = \sum_{i=1}^l l_r\left(f; p^{(t)}, q^{(\theta)}, z^{(b)}\right) + \sum_{\theta \in \Theta} \lambda_\phi \theta^2. \quad (8)$$

Where λ, θ re model parameters, θ regularization parameters, and initial parameters, knowledge point feature extraction based on the above algorithm can better comb and display the context.

2.2. Characteristic Identification of Knowledge Points in Personalized Nursing Teaching.

The purpose of personalized nursing teaching knowledge point recommendation based on the SOM neural network and ranking factor decomposition machine model is to accurately cluster users by using the nonparametric characteristics and high accuracy of the SOM neural network, and then use the characteristics of the factor decomposition machine that can easily integrate high-dimensional data as a ranking function to evaluate the academic background, quality, and accuracy of users in the same cluster. Borrow a variety of characteristic information, such as nursing teaching introduction text and web access behavior to model, and use the level sorting learning algorithm to train the model so as to realize the accurate sorting and recommendation of nursing teaching knowledge points. The flow chart of sorting recommendations based on SOM and RFM is shown in Figure 2.

Step 1. initialize the network, that is, set the SOM network and initialize the initial value of each training parameter [10, 11]. The values to be initialized are: the random number that gives the link weight W in the [0,1] interval; determine the initial value of the learning rate $n(1)\eta(o)$ $(0 < \eta(\omega) < 1)$; and determine the initial value $n(o)$ of the domain $n(t)$, where G is the winning neuron; calculate the Euclid distance between the weight vector $w = (w, m)$ and the input sample $x = (x_1, x_2, \ldots, x_n)$, and select the minimum distance to determine the winning neuron [12]. Adjust the connection weight w and update the neighborhood $n(i)$ of the output layer. The update formula for the connection weight between each neuron in the input layer and the neuron in the input layer is as follows:

$$\begin{cases} w_{ij}(t+1) = w_{ij}(t) + \eta(t) \times \left(x_j - w_{ij}(t)\right), \\ gw_{ij}(t+1) = w_{ij}(T) + \eta(t)/2 \times \left(x_j - w_{ij}(t)\right), j \in N_g(t), \\ w_{ij}(t+1) = w_{ij}(t), j \notin N_g(t). \end{cases}$$
$$(9)$$

Where $w_{ij}(t+1)$ represents the input neuron at $t+1$ time; connection weight with output neuron J; $n(t)$ is the domain range centered on the winning neuron g at time t; update the learning rate M and domain $n(t)$; then,

$$\begin{cases} \eta(t) = \eta(0) \times (1 - t/T), \\ N_g(t) = N_g(0) \times (1 - t/T). \end{cases} \quad (10)$$

The knowledge points have an inevitable sequence in the learning process. Whether a knowledge point can be learned at present often depends on whether other knowledge points have been learned, or if the latter is the preparatory knowledge of the former [13]. Before learning a certain knowledge point, you must first learn another related knowledge point, and the relationship between the two is the precursor relationship [13]. After learning a certain knowledge point, the knowledge points directly supported by this knowledge point form a successor relationship between the two, as shown in Figure 3.

As a knowledge system, there is an inherent relationship of mutual restriction and mutual influence between concepts and principles [14–16]. The relationship reveals that there is a network structure between knowledge points and points out that knowledge is composed of a group of interconnected and interactive nodes. A correlation is conducive to the mastery of knowledge and the formation of knowledge system. The association relationship between knowledge points can be divided into two categories: one-to-one association (1 : 1), which means that one knowledge point only corresponds to another knowledge point; one to many association (1 : m), which indicates that a knowledge point can be associated with multiple knowledge points; and the graph indicates the one to many association between knowledge points. Figure 4 shows the characteristics of the tree structure of nursing knowledge points.

Knowledge points are very important for teaching activities. After completing the division of knowledge points and determining the relationship between knowledge points, we should consider how to organize teaching according to knowledge points in a specific knowledge field, because teaching is composed of the teaching of knowledge points [17–19]. The knowledge point structure model can better organize and describe the content of knowledge fields. The knowledge point structure model is basically a hierarchical tree structure. In the hierarchy of relationship, the parent-child relationship and the brother relationship are the two most important relationships. They serve as the basis for building the tree structure for knowledge points. The association relationship enriches the content of the knowledge tree. Using these two relationships to describe knowledge points constitutes a knowledge point structure model. On the basis of this model, the knowledge points of a specific discipline can be listed and organized according to the relationships between knowledge points, and the knowledge point structure diagram of this discipline can be constructed. The knowledge in textbooks is generally arranged in linear order. In fact, the relationship between knowledge points is complex [20, 21]. To learn a knowledge point, you must first have certain basic knowledge (a precursor relationship), that is, master some knowledge points. To learn these knowledge points, you may need to master other knowledge points. In this way, all knowledge points and the relationship between them constitute a knowledge point network. The knowledge
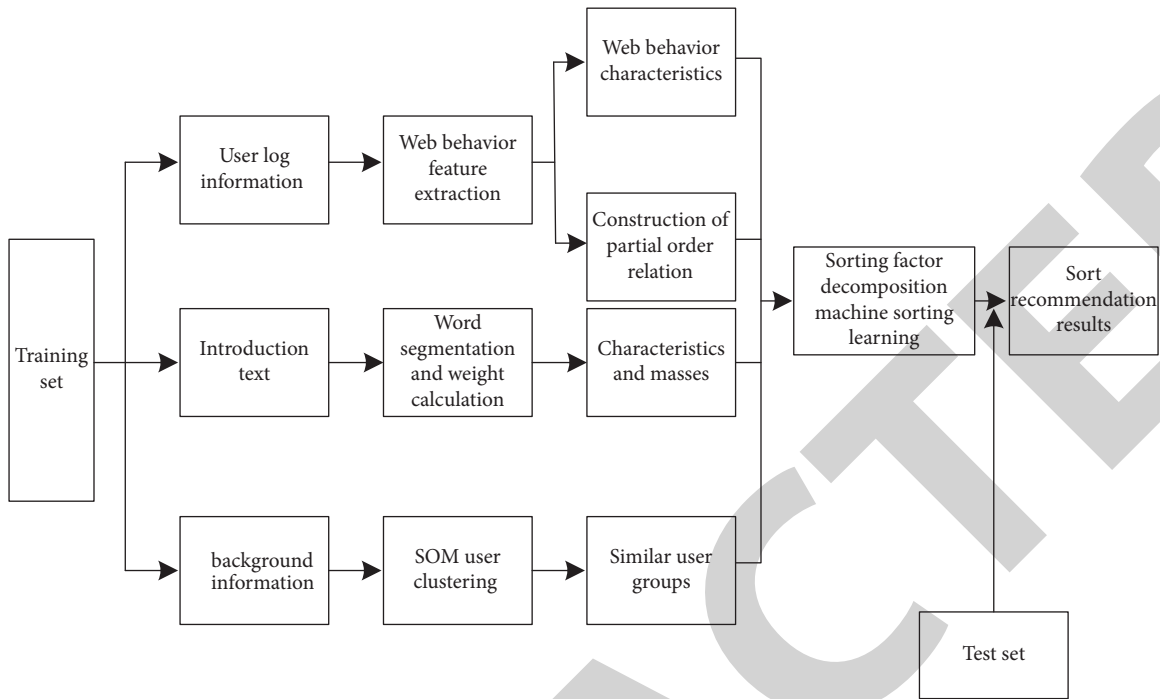
FIGURE 2: Accurate recommendation process of teaching resources based on SOM.
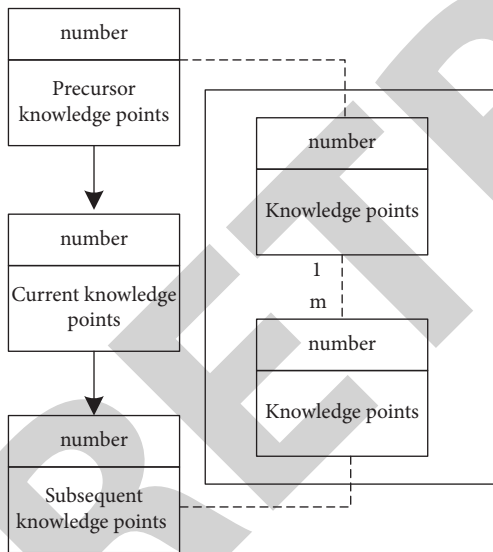
FIGURE 3: Leading and following relationships between knowledge points.

point network is a network composed of several related knowledge points based on their internal relations [19]. The nodes of the network represent knowledge points, and the links between nodes represent the links between knowledge points. After learning a knowledge point, students should also understand that the "environment" of the knowledge point is the content of its sequence, left and right, up and down, so as to determine the mark of the "status" of the knowledge, in order to make students realize the structure of the same network between the knowledge points and

establish the consciousness of the network. Through this network diagram, teachers and students can have a clear understanding of some theorems of solid geometry [22, 23]. For example, in the teaching process, if we pay attention to the application of the knowledge point network to tell students about knowledge points, it is beneficial for students to understand the knowledge structure, build their' own cognitive systems, and facilitate the transfer of memory and knowledge skills, because the knowledge point network contains information about the learning path , and this learning path should be reasonable and optimal for students. In a word, this paper deeply analyzes the relevant contents of knowledge point representation and establishes a knowledge point model that is suitable for teaching in form, reflects the connotation of knowledge points in content, and helps to realize the teaching process, which provides a new perspective for teachers to design teaching according to the attributes and laws of knowledge points.

2.3. Realization of Knowledge Point Context Visualization. The construction of a knowledge map is generally divided into two ways: top-down and bottom-up. The top-down construction method is based on ontology and takes highly structured encyclopedias and other websites as data sources to extract ontology and rule constraints and fill them into the knowledge base.

Figure 5 shows the technical architecture of the knowledge map. The three steps of knowledge extraction, fusion, and processing in the box are the core of the construction of the knowledge map. It can be seen from the figure that structured data can easily extract knowledge from it because of its high degree of standardization;
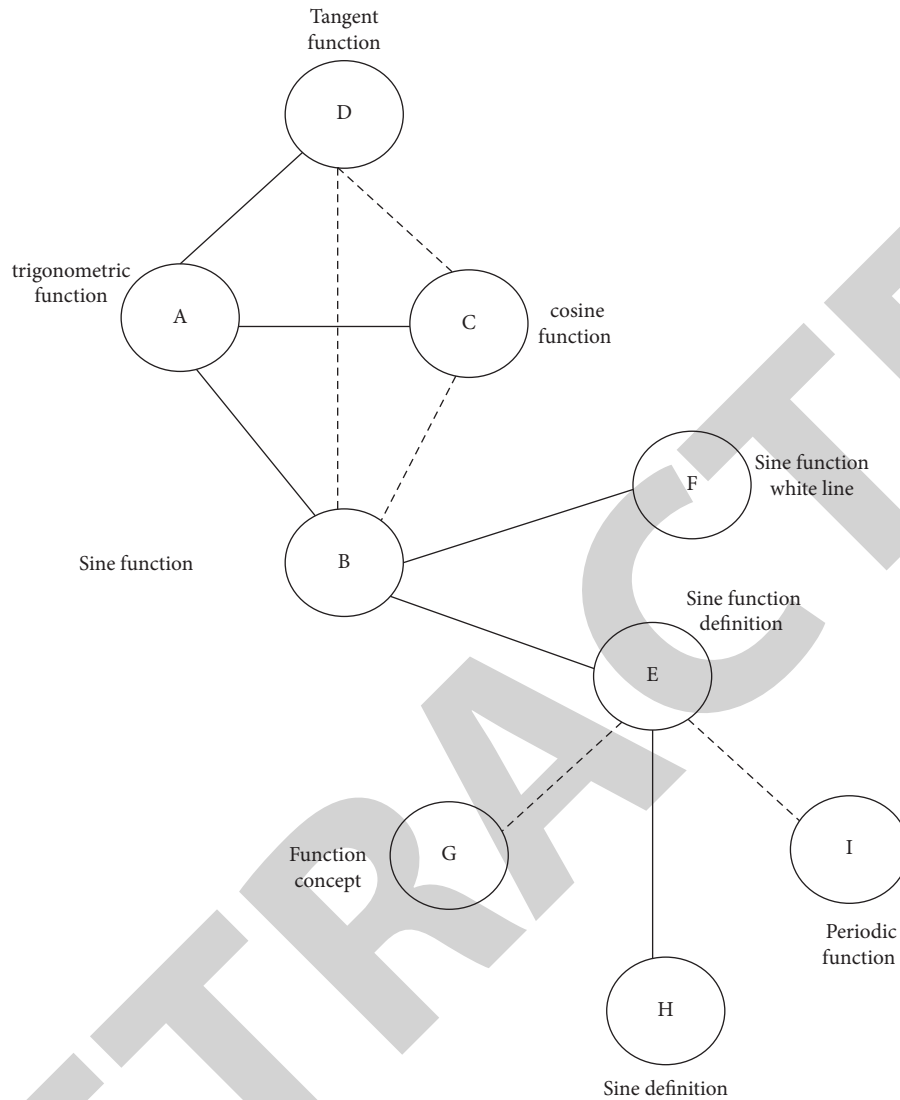
Figure 4: Tree structure characteristics of nursing knowledge points.

semistructured and unstructured data have poor standardization and are difficult to obtain knowledge directly. Therefore, it is necessary to extract the entities and associations of knowledge with the help of a series of operations such as attribute extraction, relationship extraction, and entity extraction, and then store them in the knowledge base. The construction process of a knowledge map is a continuous cycle. The iterative process can be roughly divided into three stages: knowledge extraction, knowledge fusion, and knowledge processing.

In the traditional teaching of nursing management, the teaching goal is above all else. It is not only the starting point of the teaching process but also the destination of the teaching process. However, in the classroom of network teaching, because it emphasizes that students are cognitive subjects and active constructors of meaning, students' meaning construction of knowledge is regarded as the ultimate goal of the whole learning process. The whole teaching process starts with a situation conducive to students' meaning construction and closely surrounds the

center of "meaning construction." Whether it is students' independent exploration, cooperative learning, or teachers' guidance, in short, all aspects of the learning process should belong to this center, which should be conducive to completing and deepening the meaning construction of the learned knowledge. Combined with the characteristics of nursing management and based on constructivism theory, this paper constructs the structure of nursing management teaching modes in a network environment. The specific operation flow is shown in Figure 6.

Acquiring knowledge through a well-structured relational database or third-party knowledge base is also a good choice to build a knowledge map.

Merge the ontology in the third-party knowledge base into its own library. As another important knowledge source for knowledge mapping, relational databases can usually use the resource description framework (RDF) as a data model and integrate it into a knowledge map. At present, a considerable number of open-source tools support the transformation of data in structured relational databases into
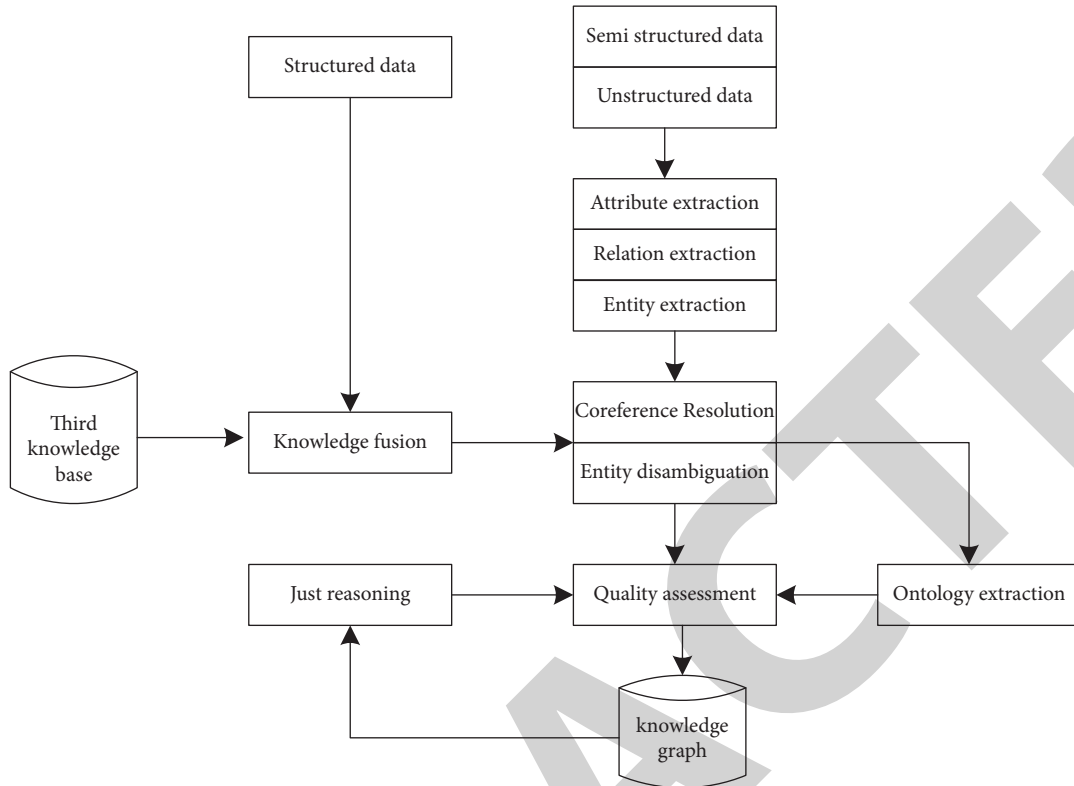
Figure 5: Structure diagram of a knowledge map visualization technology.
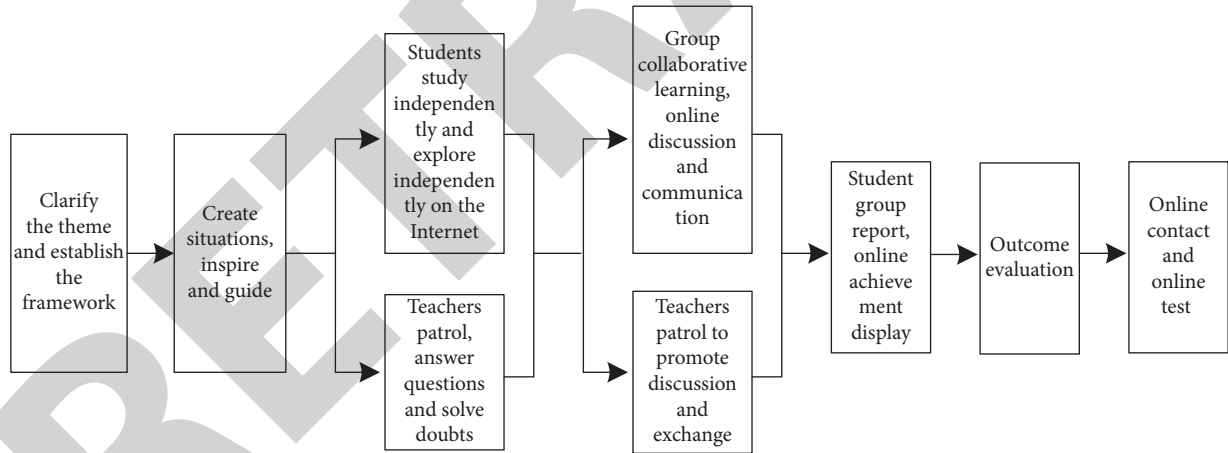
Figure 6: Basic flow chart of nursing teaching view processing under the SOM network environment.

RDF triples to realize the construction of a knowledge map. This view is a local view, which is only used to show the association between the outcome entity and the interaction entity in a specific domain. In the above view, it can be seen that for a certain nursing symptom, the associated nursing measures are very concentrated in some fields, and the visual recommendation of nursing measures can be made for the nursing measures in the same field. Therefore, we hope to design a view that can not only provide a more detailed expression of information, but also reflect the hierarchical information of nursing measures and their fields. The package layout view has the ability for hierarchical

expression and can classify and display data according to categories. However, the package view is not suitable for expressing network class information. Atlas data is a kind of network data. Network data can be expressed by force guidance diagrams, radar diagrams, chord diagrams, etc. Among them, the force guidance diagram is a node connection diagram, and the package diagram is a content filling diagram. If the two are used as a mixed view, they can achieve a complementary effect visually. Therefore, consider combining the two views. At the view level, the system interface is mainly divided into three parts. On the right is a general introduction to the Atlas data, which is divided into

the data source, data description, overall data analysis, and node selection details. The lower side is the system toolbar, which switches the interaction mode of the system. The middle part is the data view, which is used for data display and data view interaction.

3. Analysis of Experimental Results

In order to verify the effectiveness of the proposed SOM method, comparative experiments are carried out. FM is a traditional factorization machine model, which is used to judge that SOM based on ranking learning has higher accuracy than traditional FM; BPRMF is a matrix decomposition model based on pairwise ranking method, which is used to judge the influence of ranking learning algorithms on recommendation accuracy; and RSVM is a support vector machine based on pairwise ranking method, which is used to judge that FM, as a ranking function, can more accurately express user preferences than SVM, as shown in Table 1.

For ranking recommendation, because users pay more attention to the recommendation quality of the top-ranked items in the recommendation list, this study selects two ranking position-sensitive evaluation indicators for evaluation, namely average accuracy and normalized impairment cumulative gain. Map is defined as follows:

$$MAP = \frac{\sum_{r}^{N}(N_r/r \times l(r))}{N_a}. \tag{11}$$

Where r is the sorting ordinal number; N is the number of recommended products; NR is the number of related commodities sorted as r; n_r/r is the accuracy of truncated sorting; $l(r)$ is a binary correlation function with a sorting number r; correlation is 1 and uncorrelation is 0; Na is the total number of related commodities. The larger the map value, the higher the ranking of items related to user preferences, and the better the overall ranking effect of the algorithm. N_{dcg} is defined as follows:

$$NDCG@P = \frac{1}{Z_P}\sum_{i=1}^{p}\frac{2^{k(i)}-1}{lg(1+i)}. \tag{12}$$

Where P represents the position of the item in the list, Z_p is the normalization factor, and $K(i)$ represents the correlation level between the item with location i and user preference. The value range of NDCG is [0,1]. The larger the value, the more consistent the sorting results are with the user's interests and preferences. Their entity information comes from the nursing guide. By grasping the nursing guide, we extract different sets of nursing entities and the contact edge sets between entities and construct the entity network in the nursing field. The extracted main entity

Table 1: System test environment.

name	To configure
Processor	Intel(R)Core(TM) i7-7700HQ
Hard disk	1024 GB
Memory	32G
Operating system	Windows 7
Web server	Flask
Database	MySQL

information and the association information between some entities are shown in Table 2.

The traditional algorithm cannot be directly applied to the research object of this paper. In order to verify the effectiveness of the improved algorithm, this section will compare the clustering quality of the improved k-medoids algorithm with that of the improved SOM algorithm. The interclass distance and intraclass distance of the SOM improved algorithm and the k-medoids algorithm are shown in the following Table 3. In the experiment, the two algorithms are clustered ten times, and the average values of SSE and SSB from multiple experiments are calculated. As shown in Table 4, the experimental results of the two algorithms are compared.

From the perspective of stability, the repetition rates in the clustering results of the SOM improved algorithm and the k-medoids improved algorithm are compared. Here, eleven experiments are also carried out on different algorithms to calculate the average repetition rate of different algorithms.

In many experiments and simulations, the clustering results of the SOM improved algorithm can obtain a 67.0% repetition rate, which is similar to the clustering effect of the k-medoids improved algorithm, which proves that the SOM improved algorithm in this paper also has good stability. The following figure shows the proposed SOM algorithm and the comparison algorithm MAP@10 and NDCG@10 results on evaluation indicators.

It can be seen from Figure 7 that the performance of the algorithm varies with the value of K. When $k = 15$, the performance of the four algorithms is the best. At the same time, it can be seen that SOM obtains the best performance under different K values. This is because, compared with FM, SOM adopts a pairwise sorting learning algorithm, so its performance is better than FM; compared with BPRMF, SOM can not only integrate the explicit and implicit feedback information of users but also integrate the text information of the borrowing nursing teaching introduction and the borrowing log information, so its performance is also better than the traditional methods. At the same time, SOM's sorting function FM uses interaction parameters rather than independent parameters to model the interaction between features, so SOM can obtain better performance. Especially in the case of sparse data, its performance is better.

TABLE 2: Data information of nursing entities.

Number of entities in nursing symptom area	Number of nursing symptom entities	Number of nursing measures symptom entities	Number of nursing measures entities	Number of symptom domain associations	Number of symptoms associated with measures
8	130	8	410	128	2226

TABLE 3: Comparison of effectiveness between the SOM improved algorithm and the k-medoids improved algorithm.

	SSE	SSB
SOM improved algorithm	9102	13024
Improved k-medoids algorithm	9325	12058

TABLE 4: Comparison of stability between the improved algorithm of the SOM and the improved algorithm of k-medoids.

	Repetition rate (%)
SOM improved algorithm	67.0
Improved k-medoids algorithm	66.6

(a) map value

(b) ndgc value
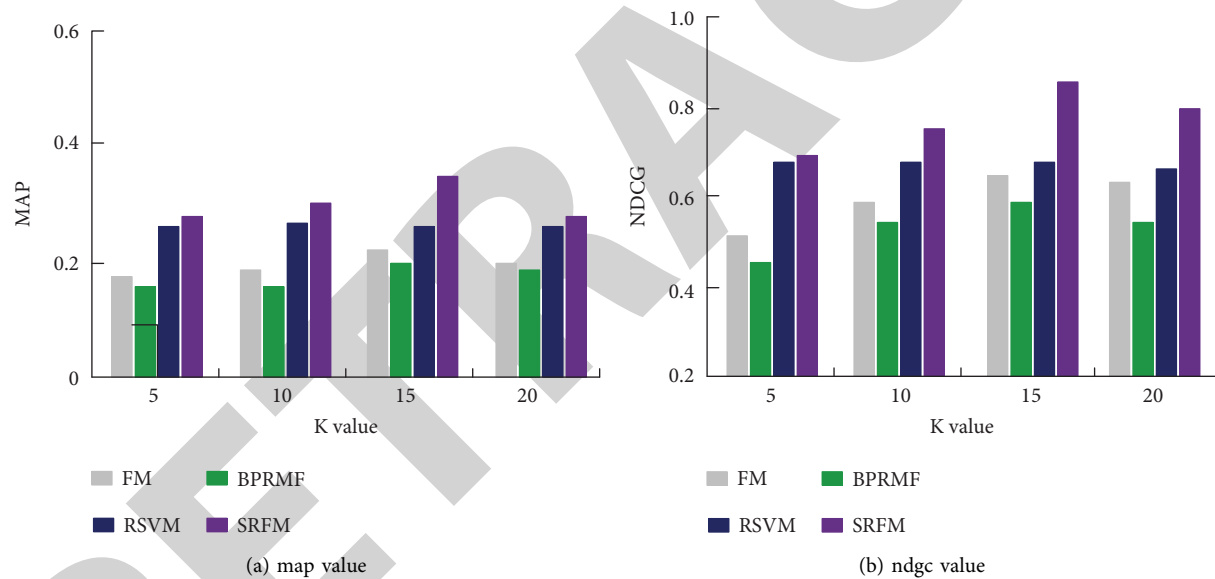
FIGURE 7: Comparison of knowledge point context view and useable effect.

4. Conclusion

In the courses that need a lot of practice, the implementation of a project-based teaching method can quickly improve students' practical operation abilities. At the same time, the establishment of students' theoretical knowledge systems cannot be ignored. The nursing specialty is a highly practical specialty. It adheres to the project-based teaching method without making students with weak theoretical basic knowledge more backward. Cultivating practical talents is one of the ways to realize this. The cultivation of operation skills is an important task in nursing teaching. Students' nursing operation level directly affects the effects of clinical practice and the future development of nursing specialties, as well as the quality of their practical talents. Years of nursing practice show that nursing teaching management is a systematic, phased, complex, and carefully organized process. In order to ensure the effect of nursing operation teaching, this paper uses a SOM neural network to cluster users according to the users' academic background information, analyzes the explicit and implicit web access behavior of nursing teaching knowledge points, constructs the partial order relationship of nursing teaching knowledge points, classifies users' academic background web access behavior, and uses a point-to-point sorting learning algorithm to accurately recommend nursing teaching knowledge points. Different teaching purposes can be implemented for each type of students. Students with weak theoretical knowledge can purposefully integrate, so that excellent students can help weak students learn from each other,

Hindawi

Retraction

Retracted: Analysis of Structured Data in Biomedicine Using Soft Computing Techniques and Computational Analysis

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. H. Rahman, "Analysis of Structured Data in Biomedicine Using Soft Computing Techniques and Computational Analysis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4711244, 11 pages, 2022.

Research Article

Analysis of Structured Data in Biomedicine Using Soft Computing Techniques and Computational Analysis

Yanping Wu[1] and Md. Habibur Rahman ![ORCID][2]

[1]Hangzhou Medical College, Hangzhou 311399, China
[2]Department of Information and Communication Technology,
 Bangabandhu Sheikh Mujibur Rahman Digital University Bangladesh, Gazipur 1750, Bangladesh

Correspondence should be addressed to Md. Habibur Rahman; habibur@ict.bdu.ac.bd

In the field of biomedicine, enormous data are generated in a structured and unstructured form every day. Soft computing techniques play a major role in the interpretation and classification of the data to make appropriate decisions for making policies. The field of medical science and biomedicine needs efficient soft computing-based methods which can process all kind of data such as structured data, categorical data, and unstructured data to generate meaningful outcome for decision-making. The soft-computing methods allow clustering of similar data, classification of data, predictions from big-data analysis, and decision-making on the basis of analysis of data. A novel method is proposed in the paper using soft-computing methods where clustering mechanisms and classification mechanisms are used to process the biomedicine data for productive outcomes. Fuzzy logic and C-means clustering are devised as a collaborative approach to analyze the biomedicine data by reducing the time and space complexity of the clustering solutions. This research work is considering categorical data, numeric data, and structured data for the interpretation of data to make further decisions. Timely decisions are very important especially in the field of biomedicine because human health and human lives are involved in this field and delays in decision-making may cause threats to human lives. The COVID-19 situation was a recent example where timely diagnosis and interpretations played significant roles in saving the lives of people. Therefore, this research work has attempted to use soft computing techniques for the successful clustering of similar medical data and for quicker interpretation of data to support the decision-making processes related to medical fields.

1. Introduction

1.1. Background. Data mining is to process where hidden information is retrieved from complex data sets by interpreting the data in an appropriate way [1]. With the increasing use of digital technology in biomedical applications, such as electronic biomedical records and digital imaging technologies, large volumes of biomedical data are collected daily [2]. The plenty of stored biomedical data has led to an urgent request for new methods and tools to transform the accumulated data into readable biomedical information [3]. This has opened up exciting opportunities for the application of data mining techniques to develop new prediction and diagnosis models in biomedicine. Clustering is an important technique in data mining to partition a set of data objects into subsets according to data similarity [4]. Similar data are

placed in one cluster. The clustering technique has been applied widely in biomedicine [5]. There are several complexities of biomedical data that challenge the application of clustering techniques in biomedicine. The first challenge is the ambiguity of biomedical data caused by the fact that one feature might be an indicator for two or more clusters with similar attributes or that certain features are not explicitly recorded in biomedical records [6]. This ambiguity can cause an overlap of cluster boundaries, i.e., the same class of feature may belong to several biomedicine clusters. In the traditional clustering technique, a hard-clustering method is used to arbitrarily partition a class of features into one cluster. Although this hard clustering method eliminates ambiguity, it can also cause information loss of that feature as an indicator for another cluster [7]. Therefore, the traditional clustering technique is suboptimal and there is a

need for newer techniques. The fuzzy clustering algorithm (FCM), a specific type of clustering technique, has been introduced to resolve this challenge, arising from the ambiguity of biomedical datum, by partitioning data according to fuzzy membership values. Fuzzy membership is the data's degree of similarity with each cluster. As one piece of data may have varying degrees of similarity with several clusters, a fuzzy clustering algorithm uses iterative algorithms, such as objective function, to compute the membership of the data into several relevant clusters with varying membership values. The FCM clustering algorithm is a specific type of fuzzy clustering algorithm developed by Bezdek [5]. For objective function, the FCM algorithm assigns a fuzzy membership value between 0 and 1 to each data object according to the data's distance to a cluster center. The notion of "fuzzy" implies that a data's membership is not fixed, but dynamic in association with its varying distance to different clusters. For example, as the symptom of chest pain can be an indicator for both congestive heart failure and chest injury, its membership to the congestive heart failure cluster can be 0.5, and to the heart failure cluster can be 0.3. Only after the possibility of chest injury is completely eliminated according to X-ray scanning results can the membership of this symptom be entirely associated with congestive heart failure. Due to its ability to handle the "fuzzy" nature of biomedical data, the FCM algorithm has been widely applied in biomedicine for such things as disease detection, biomedical image segmentation, and biomedical feature selection [8]. To date, the biggest challenge for the FCM algorithm remains how to handle the complexity of biomedical data including both categorical data (e.g., name, group) and numeric data (e.g., time, length) [9]. This shows a more realistic and accurate approach for clustering the medical data.

1.2. Literature Review. In reference [10], the authors state that different representations, processing, and computation are required for numeric and categorical data. The traditional method of using an objective function to directly measure the Euclidean distance for clustering numeric data would not work for categorical data, which does not have explicit Euclidean distance. To deal with this challenge, two methods have been widely used to convert different types of biomedical data. One is to convert the categorical data into numerical data using a binary coding technique, i.e., coding "yes" as "1," and "no" as "0." Another method is to convert numeric data into categorical data by a discretization method. For example, this might convert length into three groups, those aged between "20 to 30" as "1," "30 to 40" as "2," "40 to 50" as "3," and so on. It is no surprise that with a process as complicated as data conversion, problems emerge throughout the process. One of the common problems with data conversion is losing information. For example, if converting length "20 to 30" to "1," the useful information may be lost that the characteristics for the length group of 20 to 25 are significantly different from that of the 25 to 30 group. In reference [11], the researchers propose an algorithm to directly cluster categorical biomedical data. For

example, Huang et al. designed a k-prototypes algorithm by combining k-means and k-modes algorithms to directly cluster categorical data without conversion. In reference [12], the authors assess different clustering techniques. Data clustering techniques can be of two types: partitioning procedures and hierarchical procedures. The first one creates a hierarchy of clusters. The results are shown as a dendrogram. Partitioning method-clustering makes various partitions of objects and evaluates them by some standard. In reference [13], the authors provide biomedical researchers with an overview of the status quo of clustering algorithms. They also illustrate examples of biomedical applications based on cluster analysis. The research helps to select the most suitable clustering algorithm for different types of applications. In reference [14], the authors propose a novel method for clustering and analysis of complete biomedical article texts. The cosine coefficient is used on a subspace of two vectors. The Euclidean distance is not computed for all vectors. Then, a strategy and algorithm are introduced for semi-supervised affinity propagation (SSAP). The results show that the SSAP outperforms conventional k-means methods. In reference [15], A similarity-based agglomerative clustering (SBAC) algorithm is used to cluster data by similarity to reach a cohesive hierarchical cluster. In reference [16], a modified FCM clustering technique with a hybrid fuzzy time series model is used to deal with disease interval information to predict the infected cases and deaths of COVID-19. Some authors put forward the Kullback–Leibler FCM algorithm to process Gauss-multinomial-distributed data sets (KL-FCM-GM) [17]. In reference [18], the authors have applied a firefly algorithm in the objective function for cluster center selection. By combining the first step of cluster center selection with the second step of calculating the objective function, this algorithm effectively overcomes the limitation of local cluster optimization, and thus was useful for clustering a large medical data set [19].

1.3. Contribution of the Paper

(i) The Fuzzy C-means (FCM) clustering algorithm has been proposed to analyze biomedical datasets in this paper by adding a weighted mechanism. This study proposes a novel multiple weighted Fuzzy C-means for Mixed Data (MD-MWFCM) clustering algorithm.

(ii) The MD-MWFCM algorithm presents a novel clustering FCM algorithm by treating numeric and categorical data individually with their respective cluster center representation, dissimilarity measurement, and objective function aspects.

(iii) The model is proposed that will take care of numerical and categorical data.

(iv) The performance of the MD-WFCM algorithm has illustrated its useful application not only in a pure attribute dataset (numerical data or categorical data) but also, more importantly, in a mixed biomedical dataset.

(v) The MD-MWFCM algorithm is improved from cluster center initialization through the minimum threshold method.

2. Proposed Methodology

2.1. FCM Algorithm. There are three common data processing steps in applying the FCM algorithm:

Step 1: initializing the cluster centers and fuzzy membership matrix.

Step 2: calculating the objective function and update the cluster centers.

Step 3: continuously iterating through Steps 1 and 2 until the defined threshold of the membership cluster is reached.

One of the limitations of the FCM algorithm was, that it is oversensitive to the location of cluster centers initialized in Step 1. The selection of a cluster center is vital because the center is used by the objective function to calculate the Euclidian distance of a data point. Previous researchers have developed various methods to optimize the selection of cluster centers. Another limitation of the FCM algorithm is no discretion of the varying weights of an attribute in different clusters in Step 2. This may cause limitations for the FCM to reach an appropriate level of performance in clustering biomedical data. A data attribute was assumed to have the same weight in different clusters, which is not always the case for biomedical data. This leads to the same dissimilarity measurement used to process both numeric data and categorical data without considering the differences between the two data sets, which may cause the loss of biomedical information. The optimal algorithm would be to analyze the data in its original data format.

To improve this method, some researchers such as Xiao et al. have tried proposing a Gaussian smoothed and weighted FCM clustering algorithm (WGFCM) in brain magnetic resonance image segmentation [19]. Improvements have been made on certain aspects of FCM, i.e., optimization of initial cluster centers and assignment of weights of an attribute to various cluster centers [20]. Smart technologies have also implemented several information extraction methods in artificial intelligence, internet of things, cyber-physical system, cybersecurity and so on [21]. One of the data mining algorithms quantum adaption cuckoo search (QACS) is used to identify the unauthorised users by extracting the important features in blockchain technology [22]. However, there is a need for an effective method and process in clustering mixed numerical and categorical data for the biomedical data set. To address this challenge, we propose a novel multiple weighted Fuzzy C-means for mixed data (MD-MWFCM). Instead of data transformation, the MD-MWFCM will improve data representation by directly using the original data type for clustering analysis. It will use different dissimilarity measurement methods to treat different data types and then calculate the distance of a data point with weighting to various cluster centers. Different methods will be used to represent cluster centers for different data types; mean will be used for numeric data and fuzzy center will be used for categorical data. This will lead to improvement in the selection of the initial cluster center.

2.2. The MD-MWFCM Algorithm. k-Means is a widely used algorithm for clustering. Its limitation is that it only works for numerical data. It is not suitable for categorical data types. Modified variation is k-Modes which is created in order to handle clustering algorithms with the categorical data type. The limitation of k-mode is that it can handle categorical data only. The proposed algorithm can handle numerical as well as categorical data. Mean will be used for numeric data and fuzzy center will be used for categorical data. This will lead to improvement in the selection of the initial cluster center.

The proposed method also introduces a cluster initialization strategy in which the attributes have been assigned weights. The attribute domains include both numeric and categorical domains. A new method is proposed to measure the similarity level between the value of a categorical attribute of the variable to the center of a categorical cluster . It is based on a method proposed in the article [7]. Binary distance $\delta(x\_ij, a\_kjt)$ is used to measure the similarity between the variable and cluster center. The resulting value of $\delta(x\_ij, a\_kjt)$ is between 0 and 1 depending on the level of similarity between the variable and cluster center. This will lead to improvement in the selection of the initial cluster center.

2.2.1. Data Notation. Let $X = \{X_1, X_2, \cdots, X_n\}$ denote a set of n data points to be clustered and $X_i (i = 1, 2, \cdots, n)$ be a data point with m attributes $\{A_1, A_2, \cdots, A_m\}$. Each attribute $A_j (j = 1, 2, \cdots, m)$ includes a domain of values denoted by $\text{Dom}(A)_j$. A data point X_i can be expressed with attributes as in equation (1) as follows:

$$[A_1 = x_{i1}] \wedge [A_2 = x_{i2}] \wedge \cdots \wedge [A_j = x_{ij}] \cdots \wedge [A_m = x_{im}],$$

$$(1)$$

where $x_{ij} \in \text{Dom}(A_j)$ for $j = 1, 2, \cdots, m$. Values to the set of attribute weights are given in equation (2) as follows:

$$W = [w_1, w_2, \cdots, w_j, \cdots, w_m]^T,$$
$$\forall j: w_j > 0, \sum_{j=1}^{m} w_j = 1.$$
$$(2)$$

The attribute domains include both numeric and categorical domains. The categorical domain is denoted by $\text{Dom}(A)_j = \{a_j^1, a_j^2, \cdots, a_j^{n_l}\}$, where n_l is the value number of a categorical attribute j. X_i is represented as a vector. It represents the data points to be clustered. The same is given in equation (3) as follows:

$$X_i = [x_{i1}^r, x_{i2}^r, \cdots, x_{ip}^r, x_{i(p+1)}^c, x_{i(p+2)}^c, \cdots, x_{im}^c]$$
$$(i = 1, 2, \cdots, n),$$
$$(3)$$

where elements p with superscript r and p with superscript c are numeric values and categorical values, respectively, and each $X_i (i = 1, 2, \cdots, n)$ with m attributes. The attribute weight is represented in equation (4).

$$W = [w_1^r, w_2^r, \cdots, w_p^r, w_{(p+1)}^c, w_{(p+2)}^c, \cdots, w_m^c]^T. \quad (4)$$

2.2.2. Fuzzy Centre for Categorical Data. A set of n data points $X = \{X_1, X_2, \cdots, X_n\}$ have both numeric and categorical attributes. They belong to C clusters. Their corresponding cluster center set is calculated in equation (5) as follows:

$$V = [V_1, \cdots, V_k, \cdots, V_c] \, (k = 1, 2, \cdots, C). \tag{5}$$

Each cluster center $V_k \, (k = 1, 2, \cdots, C)$, includes numeric and categorical attributes and can be represented equation (6) as follows:

$$V_k = \left[v_{i1}^r, v_{i2}^r, \cdots, v_{ip}^r, v_{i(p+1)}^c, v_{i(p+2)}^c, \cdots, v_{im}^c\right], \tag{6}$$

where elements p are defined in equation (3). Every cluster center $V_k \, (k = 1, 2, \cdots, C)$ has exactly m attributes. For each of the categorical data of the mixed data set $X_k^c = [X_{i(p+1)}^c, X_{i(p+2)}^c, \cdots, X_{im}^c]$, there exist C fuzzy centers $V_k^c = \tilde{v}_k \, (k = 1, 2, \cdots, C)$ [5]. For a fuzzy center $\tilde{v}_k \, (k = 1, 2, \cdots, C)$ in the categorical data domain $\text{Dom}(A_j) = \{a_j^1, a_j^2, \cdots, a_j^t, \cdots, a_j^{n_l}\}$, the fuzzy center of the categorical data set, denoted by $\tilde{V}$, is defined in equations (7) as follows:

$$\begin{aligned} \tilde{V} &= [\tilde{v}_1, \cdots \tilde{v}_k, \cdots, \tilde{v}_c] \, (k = 1, 2, \cdots, C), \\ \tilde{v}_k &= \left[v_{i(p+1)}^c, v_{(ip+2)}^c, \cdots, v_{im}^c\right]. \end{aligned} \tag{7}$$

The cluster center $\tilde{v}_k^c$ of the categorical attribute $j \, (j = (p+1), (p+2), \cdots, m)$ is as in equation (8):

$$v_{kj}^c = \frac{a_j^1}{\omega_{kj}^1} = +\frac{a_j^2}{\omega_{kj}^2} + \cdots + \frac{a_j^t}{\omega_{kj}^t} + \cdots + \frac{a_j^{n_l}}{\omega_{kj}^{n_l}}, \tag{8}$$

where ω_{kj}^t is subject to the condition given in equation (9).

$$0 \leq \omega_{kj}^t \leq 1, \sum_{t=1}^{n_l} \omega_{kj}^t = 1, (j = (p+1), (p+2), \cdots, m). \tag{9}$$

The cluster center $\tilde{v}_k^c$ of the categorical attribute j can be deduced in equation (11) as follows

$$v_{kj}^c = \sum_{t=1}^{n_l} \frac{\sum_{i=1}^n 1\left(x_{ij} = a_j^t\right)(\mu_{ik})^\alpha}{\sum_{i=1}^n (\mu_{ik})^\alpha} a_j^t \, (j = (p+1), (p+2), \cdots, m), \tag{10}$$

where n_l is the value number of a categorical attribute j, t is the $t \, (t = 1, 2, \cdots, n_l)$ value of a categorical attribute j, μ_{ik} indicates the fuzzy membership degree of data $X_i \, (i = 1, 2, \cdots, n)$ to the k th $(k = 1, 2, \cdots, C)$ cluster. μ_{ik} is an element of the partition matrix $U_{n \times c}$, and subject to condition given in equation (11) as follows-

$$\sum_{k=1}^C \mu_{ik} = 1, 0 \leq \mu_{ik} \leq 1. \tag{11}$$

Each attribute has a fuzzy categorical value $\tilde{v}_j \in \tilde{V}$, represented as a fuzzy set $\{(a_j^t, \omega_j^t)\} \, (t = 1, 2, \cdots, n_l)$. The parameter α, the fuzzy coefficient exponent, controls the fuzzy degree of partition matrix $U \in R^{C \times n}$. According to

previous experience, for numbers in clusters 2–10, the best choice for α is from 1.01 to 7.0.

2.2.3. Similarity Level between Two Categorical Attributes. Let x denote categorical data and y be the cluster center of x. The previous studies measure the similarity between x and y with function $\delta(x, y)$. This has resulted in the "hard partition" of x into only one categorical cluster y if $x = y$ or vice versa if $x \neq y$. Namely, $\delta(x, y) = \begin{cases} 0, x \neq y \\ 1, x = y \end{cases}$. A new method is proposed to measure similarity level between the value of a categorical attribute x_{ij} of the jth variable to the center of a categorical cluster a_{kj}^t based on a method proposed in article [7]. We use binary distance $\delta(x_{ij}, a_{kj}^t)$ to measure similarity between x_{ij} and $a_{kj}^t \, (1 \leq t \leq n_l)$. The resulting value of $\delta(x_{ij}, a_{kj}^t)$ is between 0 and 1 depending on the level of similarity between x_{ij} and a_{kj}^t. The similarity function is stated as follows in equation (12).

$$\delta\left(x_{ij}, a_{kj}^t\right) = \frac{1}{C}\left(\sum_{k=1}^C \left|\frac{N_{x_{ij}, k}}{N_{x_{ij}}} - \frac{N_{a_{kj}^t}}{N_{a_j^t}}\right|\right), \tag{12}$$

where $N_{x_{ij}, k}$ is the data number in dataset $X_i \, (i = 1, 2, \cdots, n)$ with the value x_{ij} for the $j^{th} \, (j = (p+1), (p+2), \cdots, m)$ attribute of the cluster $k \, (k = 1, 2, \cdots, C)$. With the fuzzy membership $(\mu_{ik})^\alpha$ of the data point x_{ij}, we can compute $N_{x_{ij}, k}$, the association of value x_{ij} for j^{th} attribute within cluster k. The data number in dataset X_i is calculated in equation (13) as follows:

$$N_{x_{ij}, k} = \sum_{l=1, l \neq i}^n 1\left(x_{lj} = x_{ij}\right)(\mu_{ik})^\alpha, \tag{13}$$

where $1\,(x_{lj} = x_{ij}) = 1$ if $x_{lj} = x_{ij}$, where $1\,(x_{lj} \neq x_{ij}) = 0$ if $x_{lj} \neq x_{ij}$. $N_{x_{ij}}$ is computed in the equation (14) as follows:

$$N_{x_{ij}} = \sum_{i=1}^n (\mu_{ik})^\alpha. \tag{14}$$

2.2.4. Attribute Weight of Categorical Data. Let $W = [w_1^r, w_2^r, \cdots, w_p^r, w_{(p+1)}^c, w_{(p+2)}^c, \cdots, w_m^c]^T$ be the weights for m attributes. Superscript c denotes weight for categorical attributes $W^C = [w_{(p+1)}^c, w_{(p+2)}^c, \cdots, w_m^c]^T$. Assuming $N_{a_j^{t_1}} \, (t_1 = 1, 2, \cdots, n_l)$ and $N_{a_j^{t_2}} \, (t_1 = 1, 2, \cdots, n_l)$ are data point numbers in Dataset $X_i \, (i = 1, 2, \cdots, n)$ with the value $a_j^{t_1}, a_j^{t_2}$ for the $j^{th} \, (j = (p+1), (p+2), \cdots, m)$ attribute that belongs to cluster $k \, (k = 1, 2, \cdots, C)$, respectively. $N_{a_j^{t_1}}$ and $N_{a_j^{t_2}}$ are data point number in Dataset X_i with $a_j^{t_1}$ and $a_j^{t_2}$ for the jth attribute in all clusters, respectively. Based on providing the number of formula such as (2) and (3), the weight of categorical attribute $A_j \, (j = (p+1), (p+2), \cdots, m)$ can be calculated by weighting distance $\vartheta(a_j^{t_1}, a_j^{t_2})$. Thus, the weight of attribute $A_j \, (j = (p+1), (p+2), \cdots, m)$ in m categorical attributes is as shown in the equation (15) as follows-

$$w_j^c = \frac{\sum_{t_1,t_2=1}^{n_l} \vartheta\left(a_j^{t_1}, a_j^{t_2}\right)}{n_l}. \tag{15}$$

The weighted distance between $a_j^{t_1}$ and $a_j^{t_2}$ is given in equation (16) as follows:

$$\vartheta\left(a_j^{t_1}, a_j^{t_2}\right) = \sum_{k=1}^{C} \left| \frac{N_{a_{kj}^{t_1}}}{N_{a_j^{t_1}}} - \frac{N_{a_{kj}^{t_2}}}{N_{a_j^{t_2}}} \right|. \tag{16}$$

2.2.5. Fuzzy Centre for Categorical Data Objective Function.

We will apply the common objective function in the FCM algorithm [5], as presented in equation (17) as follows:

$$J = \sum_{k=1}^{C} \sum_{i=1}^{n} (\mu_{ik})^\alpha d(X_i, V_k). \tag{17}$$

When a dataset only has numeric attributes, the similarity is measured by the square of Euclidean distance. When a data set has mixed attributes of numerical and categorical data types, the similarity measure needs to incorporate weights of attributes to various cluster centers. Assuming a dataset $X = \{X_1, X_2, \cdots, X_n\}$ has both numerical and categorical attributes. They belong to C clusters. Partition matrix $U_{n \times c}$ indicates the fuzzy membership degree of data i to the kth cluster. μ_{ik} is an element in the partition matrix $U_{n \times c}$ and subject to condition given in equation (18) as follows-

$$\sum_{k=1}^{C} \mu_{ik} = 1, 0 \le \mu_{ik} \le 1. \tag{18}$$

Their corresponding cluster center set is as given in equation (19) as follows-

$$V = [V_1, \cdots, V_k, \cdots, V_c]. \tag{19}$$

Each cluster center $V_k (k = 1, 2, \cdots, C)$, includes numerical and categorical attributes and can be represented as in equation (20)

$$V_k = \left[v_{i1}^r, v_{i2}^r, \cdots, v_{ip}^r, v_{i(p+1)}^c, v_{i(p+2)}^c, \cdots, v_{im}^c\right]. \tag{20}$$

The weights for m attributes is given in equation (21):

$$W = \left[w_1^r, w_2^r, \cdots, w_p^r, w_{(p+1)}^c, w_{(p+2)}^c, \cdots, w_m^c\right]^T, \tag{21}$$

where elements p with superscript r and p with superscript c are numeric values and categorical values respectively. The common objective function is given in equation (22)

$$J(U, V, W) = \sum_{k=1}^{C} \sum_{i=1}^{n} \mu_{ik}^\alpha d_w^2(X_i, V_k) = \sum_{k=1}^{C} \sum_{i=1}^{n} \mu_{ik}^\alpha \|X_i - V_k\|_w^2. \tag{22}$$

The common objective function is represented in equation (23) as follows:

$$J(U, V, W) = J^r(U, V, W) + J^c(U, V, W). \tag{23}$$

Proof. $J(U, V, W) = \sum_{k=1}^{C}\{\mu_{ik}^\alpha [\sum_{i=1}^{n} [W\|X_i - V_k\|]^2]\} = \sum_{k=1}^{C}\{\mu_{ik}^\alpha [\sum_{i=1}^{n} \sum_{j=0}^{m} [w_j\|x_{ij} - v_{kj}\|]^2]\} = \sum_{k=1}^{C}\{\mu_{ik}^\alpha [\sum_{i=1}^{n} (\sum_{j=0}^{p} (w_j\|x_{ij} - v_{kj}\|)^2 + \sum_{j=p+1}^{m} (w_j\|x_{ij} - v_{kj}\|)^2)]\} = \sum_{k=1}^{C}\{\mu_{ik}^\alpha [\sum_{i=1}^{n} \sum_{j=0}^{p} (w_j\|x_{ij} - v_{kj}\|)^2 + \sum_{i=0}^{n} \sum_{j=p+1}^{m} (w_j\|x_{ij} - v_{kj}\|)^2]\} = \sum_{k=1}^{C}\{\mu_{ik}^\alpha [\sum_{i=1}^{n} (w^r\|X_i^r - V_k^r\|)^2 + \sum_{i=1}^{n} (w^c\|X_i^c - V_k^c\|)^2]\} = \sum_{k=1}^{C} \sum_{i=1}^{n} \mu_{ik}^\alpha (w^r\|X_i^r - V_k^r\|)^2 + \sum_{k=1}^{C} \sum_{i=1}^{n} \mu_{ik}^\alpha (w^c\|X_i^c - V_k^c\|)^2 = J^r(U, V, W) + J^c(U, V, W)$

According to the above computation, the distance $d_w^2(X_i, V_k)$ is as given in equation (24) as follows-:

$$d_w^2(X_i, V_k) = d_w^2(X_i^r, V_k^r) + d_w^2(X_i^c, V_k^c),$$

$$d_w^2(X_i^r, V_k^r) = \|X_i^r - V_k^r\|_w^2 = \left(w^r\|X_i^r - V_k^r\|\right)^2$$

$$= \sum_{j=1}^{p} \left[w_j\left(x_{ij} - v_{kj}\right)\right]^2,$$

$$d_w^2(X_i^c, V_k^c) = \frac{1}{N_k} \sum_{t=1}^{n_l} N_{a_{kj}^t} \delta\left(x_{ij}^c, a_j^t\right)$$

$$= \frac{1}{N_k} \sum_{t=1}^{n_l} \left(N_{a_{kj}^t} \times \frac{1}{C} \left(\sum_{p=1}^{C} \left|\frac{N_{x_{ij},k}}{N_{x_{ij}}} - \frac{N_{a_{kj}^t}}{N_{a_j^t}}\right|\right)\right). \tag{24}$$

$\square$

2.2.6. The Procedure of Applying the MD-MWFCM Algorithm.

Let the dataset of biomedicine be $X = \{X_1, X_2, \cdots, X_n\}$. For a data point $X_i (i = 1, 2, \cdots, n)$ having m features $m = \{A\}_1, A_2, \cdots, A_m$, the feature data set can be represented as $X_i = \{x_{i1}, x_{i2}, \cdots, x_{im}\}$. The mixed numeric and categorical feature data set is represented as: $X_i = [x_{i1}^r, x_{i2}^r, \cdots, x_{ip}^r, x_{i(p+1)}^c, x_{i(p+2)}^c, \cdots, x_{im}^c] (i = 1, 2, \cdots, n)$; weight $W = [w_1, w_2, \cdots, w_m]^T$ represents the weight of a feature for the a biomedical clusters $C = \{C_1, C_2, \cdots, C_c\}$. This weight needs to be merged into the corresponding feature data set $W = [w_1^r, w_2^r, \cdots, w_p^r, w_{(p+1)}^c, w_{(p+2)}^c, \cdots, w_m^c]^T$. The corresponding feature fuzzy membership is represented as a matrix $U \in R^{C \times n} (\tilde{U} \in [\mu_{ik}(X_i)]_{i=1,2,\cdots,n; k=1,2,\cdots,C})$ under the following constraints given in equation (25).

$$\sum_{k=1}^{C} \mu_{ik}(X_i) = 1, 0 \le \mu_{ik}(X_i) \le 1, \tag{25}$$

where $\mu_{ik}(X_i)$ represents the fuzzy membership of biomedical data point $X_i (i = 1, 2, \cdots, n)$ to the cluster $k (k = 1, 2, \cdots, C)$. In addition to using the original formula in the FCM algorithm to compute the center of symptom j

in biomedical cluster k for numeric symptom data of dataset $X_i^r = [x_{i1}^r, \cdots, x_{ij}^r, \cdots, x_{ip}^r]$ $(i = 1, 2, \cdots, n; j = 1, 2, \cdots, p)$, the MD-MWFCM algorithm will also calculate the fuzzy membership μ_{ik} for r both numerical and categorical feature data. Each cluster center V_k can be expressed as in equation (24) and the fuzzy membership is as given in equation (26) as follows-

$$v_{kj}^r = \frac{\sum_{i=1}^n \mu_{ik}^\alpha x_{ij}^r}{\sum_{i=1}^n \mu_{ik}^\alpha} \; (j = 1, 2, \cdots, p)$$

$$\mu_{ik} = \frac{1}{\sum_{t=1}^C \left(d(X_i, V_k)/d(X_i, V_t) \right)^{1/\alpha-1}}, \quad (26)$$

where $d(X_i, V_k)$ and $d(X_i, V_t)$ is the dissimilarity measure between each biomedical data point $X_i (i = 1, 2, \cdots, n)$ and each cluster center $V_k (k = 1, 2, \cdots, C; k \neq t)$ and $V_t (k = 1, 2, \cdots, C; t \neq k)$. The parameter α is a fuzzy coefficient exponent for fuzzy degree of partition matrix $U \in R^{C \times n}$. To overcome the limitation of the FCM method in over sensitivity to initial cluster centers location, MD-MWFCM uses the minimum threshold method to initiate the cluster centers. First, it randomly locates two biomedical data points $X_x, X_y (x, y = 1, 2, \cdots, n; x \neq y)$ into a C_1. Then, it partitions the rest of the $X_i (i = 1, 2, \cdots, n; i \neq x, i \neq y)$ according to the result of the dissimilarity measure $d_w(X_i, X_x)$ or $d_w(X_i, X_y)$, where the function $d_w(X_i, X_x)$ and $d_w(X_i, X_y)$ denotes comparison of similarity between X_i and X_x or X_y. Assuming D as the threshold value for the dissimilarity measure, if $d_w(X_i, X_x) < D$ or $d_w(X_i, X_y) < D$, then we allocate X_i into the cluster C_1, otherwise allocate X_i into a new data cluster. We repeat this partition process until all biomedical data belong to certain clusters. A seven steps flowchart in implementing the MD-MWFCM algorithm is shown in Figure 1

The steps of the algorithms are as follows

Step 1. Input a normalized biomedical data $X_i = \{x_{i1}, x_{i2}, \cdots, x_{im}\} (i = 1, 2, \cdots, n)$. Select the maximum iteration number max_Iteration, a cluster number C, a sensitivity threshold value ε, a threshold value D of minimum threshold method, and fuzzy coefficient α. Initialize fuzzy membership of the feature data points $\tilde{U} \in [\mu_{ik}(X_i)]_{i=1,2,\cdots,n; k=1,2,\cdots,C}$.

Step 2. Select the initial cluster center using minimum threshold method.

Step 2.1. Randomly allocate two biomedical data points $X_x, X_y (x, y = 1, 2, \cdots, n; x \neq y)$ into a cluster C_1.

Step 2.2. Compute the dissimilarity of the rest of the biomedical data point $X_i (i = 1, 2, \cdots, n; i \neq x, i \neq y)$ with the original two data points X_x and X_y using the formulae (21)–(23). $d_w(X_i, X_x) = d_w^2(X_i^r, X_x^r) + d_w^2(X_i^c, X_x^c) d_w(X_i, X_y) = d_w^2(X_i^r, X_y^r) + d_w^2(X_i^c, X_y^c)$ If $d_w(X_i, X_x) < D$ or $d_w(X_i, X_y) < D$, allocate X_i into biomedical cluster C_1, otherwise allocate X_i into a new cluster.

Step 2.3. Repeat Step 2.2 until every data $X_i (i = 1, 2, \cdots, n)$ has been allocated into C clusters.

Step 2.4. If k = C, go to Step 2.5. If k ≠ C, change threshold value D and return to Step 2.1.

Step 2.5. Calculate the initial cluster center $V_k = [v_{i1}^r, \cdots, v_{ip}^r, v_{i(p+1)}^c, \cdots, v_{im}^c]$ $(k = 1, 2, \cdots, C)$ for each cluster using equations (24) and (11).

Step 3. Use the cluster centers $V_k (k = 1, 2, \cdots, C)$ calculated in Steps 2.5 as the initial cluster centers $V_k^{(0)} (k = 1, 2, \cdots, C)$. Define the iteration count number $l = 1$.

Step 4. Compute fuzzy membership of the feature data points $U \in R^{C \times n}$ by the equation (27), adding the iteration count number l.

$$\mu_{ik}^{(l)} = \left[\sum_{t=1}^C \left(\frac{d_w^2(X_i, V_k^{(l-1)})}{d_w^2(X_i, V_t^{(l-1)})} \right)^{\frac{1}{\alpha-1}} \right]^{-1}. \quad (27)$$

Step 5. Compute cluster centers $V_k^{(l)} = [v_{i1}^r, \cdots, v_{ip}^r, v_{i(p+1)}^c, \cdots, v_{im}^c]^{(l)} (k = 1, 2, \cdots, C)$ with equations (24) and (11) for numeric and categorical feature data.

Step 6. Compute feature weights $W^{(l)} = [w_1, w_2, \cdots, w_m]^{(l)T}$ for numeric feature data using equation (28) and categorical feature data using equation (29).

$$w_j^r = \left[\sum_{t=1}^m \frac{\sum_{k=1}^C \sum_{i=1}^n \mu_{ik}^\alpha (X_{ij} - V_{kj})^2}{\sum_{k=1}^C \sum_{i=1}^n \mu_{ik}^\alpha (X_{it} - V_{kt})^2} \right]^{-1},$$

$$(t = 1, 2, \cdots, p; j = 1, 2, \cdots, p)$$

$$w_j^c = \frac{\sum_{t_1, t_2=1}^{n_l} \vartheta(a_j^{t_1}, a_j^{t_2})}{n_l} \; (t = 1, 2, \cdots, p; j = 1, 2, \cdots, p). \quad (28)$$

Step 7. Judge objective function value $J(U, V, W)$ using formula (20). If $|J^{(l)}(U, V, W) - J^{(l-1)}(U, V, W)| > \varepsilon$ or $l < $ max _Iteration, return to Step 4 and increase $l = l + 1$. Otherwise, conclude.

3. Research Method

3.1. Data Description.
Three datasets were used in the experiment and were all exported from the UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/datasets.php). These include an Iris data set with pure numerical data, a Soybean data set with pure categorical data and a heart disease data set with mixed numeric and categorical data. The Iris data set has only four numeric attributes (sepal length, sepal width, petal length and petal width) and contains 150 data points equally in three clusters: iris Setosa (50), iris Versicolour (50), and iris Virginica (50). Since all attributes of the data set are numeric, only numeric objective function $J^r(U, V, W)$ is applied in the MD-MWFCM. The Soybean data set has all 35 categorical attributes (date, plant-stand, precip, temp, hail, crop-hist, area-damaged, severity, seed-tmt, germination, Plant-growth, leaves, leafspots-halo, leafspots-marg, leafspot-size, leaf-shread, leaf-malf, leaf-mild,
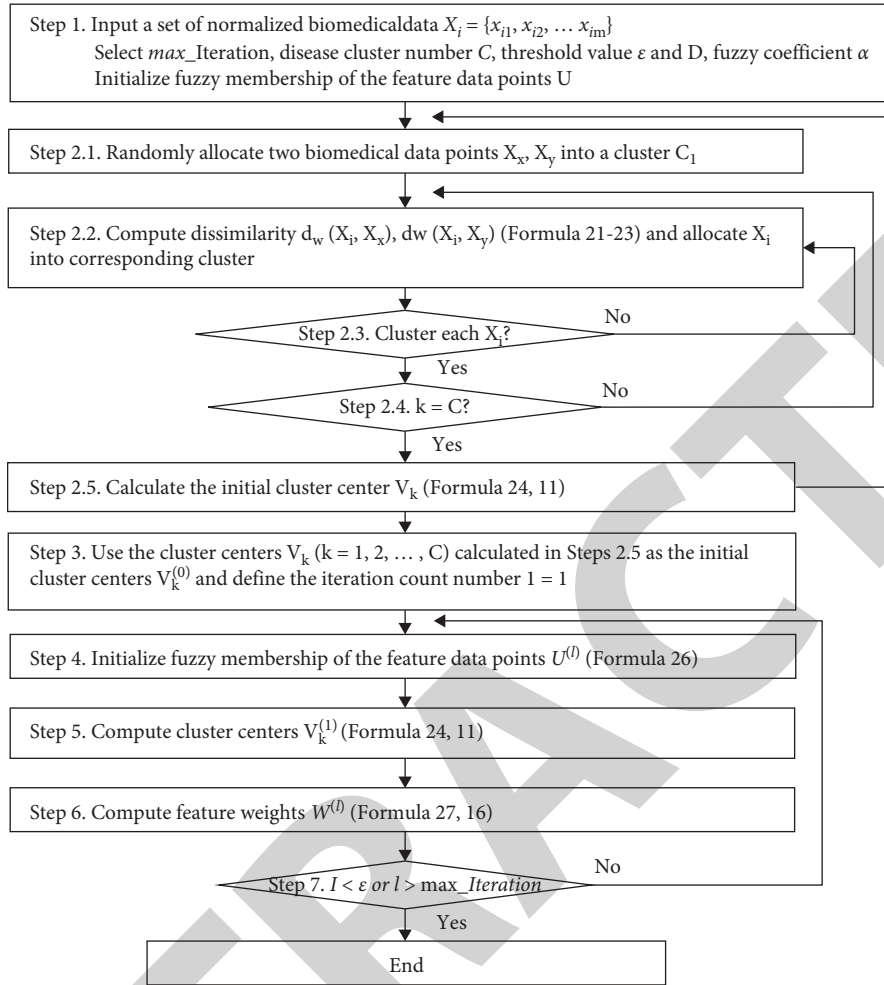
Figure 1: MD-MWFCM algorithm.

stem, lodging, stem-cankers, canker-lesion, fruiting-bodies, external decay, mycelium, int-discolor, sclerotia, fruit-pods, fruit spots, seed, mold-growth, seed-discolor, seed-size, shriveling, roots) and 47 data points. The data points are grouped into four clusters: Diapor the Stem Canker (10), Charcoal Rot (10), Rhizoctonia Root Rot (10) and Phytophthora Rot (17). Only categorical objective function $J^c(U, V, W)$ is needed to cluster all categorical data attributes.

3.2. Experiment Setup. To implement the MD-MWFCM clustering algorithm, data initialization in Step 1 included the folloiwng: setting the maximum iteration number max _Iteration = 10000, the cluster number C to be 3, 4 and 2 for Iris, Soybean, and Heart Disease data sets, respectively (see Table 1), the sensitivity threshold value $\varepsilon = 0.0001$, the threshold value D = 2 and fuzzy coefficient $\alpha \in [1.0, 10.0]$ (see Table 1). Initialize fuzzy membership values $\mu_{ik}(X_i) \in (0.0, 1.0)\,(i = 1, 2, \cdots, n; k = 1, 2, \cdots, C)$ for random function under the constraint $\sum_{k=1}^{C} \mu_{ik}(X_i) = 1$. The MD-MWFCM clustering algorithm was run for each data set. Since the three datasets are standard datasets which have been normalized in the UCI Machine Learning

Table 1: Comparison of the maximum accuracy achieved by five clustering algorithms for the Iris data set.

Algorithm	Clustering accuracy (r)
k-prototypes	0.819
SBAC	0.373
KL-FCM-GM	0.335 ($\alpha = 1.1$)
IWKM	0.822
MD-MWFCM	0.967 ($\alpha = 1.5$)

Repository, in Step 1 (see Figure 1), the data points with the relevant attributes are input parameters to run the MD-MWFCM clustering algorithm for each data set (Iris, Soybean and Heart Disease, respectively) (See Table 1). Steps 2 to 7 (Figure 1) are executed automatically by the software programs. Each dataset was run 100 times for each value of the fuzzy coefficient {1.0, 1.1, 1.2, . . . 10.0} to select the optimal true positive value as the output result at each value of the fuzzy coefficient (see Table 1). Each output result represents a cluster centre. All experiments are implemented on a Microsoft ThinkPad E450 PC with 8G byte main memory and Microsoft Windows XP operating system. The algorithms were implemented in Visual C++ programming language. Input parameters of

TABLE 2: Input parameters of each dataset for executing the MD-MWFCM clustering algorithm.

Input parameter	Iris dataset	Soybean dataset	Heart disease dataset
n (Sample size)	150	47	303
m (Attribute number)	4	35	13
C (Cluster number)	3	4	2
α (Fuzzy coefficient)	[1.0, 10.0]	[1.0, 10.0]	[1.0, 10.0]
$m - p$ (Categorical attribute number)	0	35	8
$Dom(A)$ (Categorical attribute domain)	See iris dataset	See soybean dataset	See heart disease dataset
The value of each data point		See UCI machine learning repository	

each dataset for executing the MD-MWFCM clustering algorithm are given in Table 2

3.3. Performance Measurement.
The performance of the algorithms is accessed by clustering accuracy (r) as given in equation (29)

$$r = \frac{\sum_{k=1}^{c} a_k}{n}, \tag{29}$$

where a_k is the data point number in the kth $(k = 1, 2, \cdots, C)$ cluster with true positive value. n is the total number of the data set. A higher value of r suggests better performance of the algorithm.

3.4. Result and Discussion

3.4.1. The Performance of the MD-MWFCM for Mixed Data.
The MD-MWFCM algorithm have achieved the highest clustering accuracy for three data sets compared with the four common clustering algorithms: k-prototypes, SBAC, KL-FCM-GM and IWKM. The value of r is 0.967 in clustering the numeric Iris data set by the MD-MWFCM. This is 14.8%, 59.4%, 63.2% and 14.5% higher than k-prototypes, SABC, KL-FCM-GM and IWKM algorithms (see Table 3). For the categorical Soybean dataset, *r* of the MD-MWFCM is 1, achieving clear grouping of clusters. It outperformed k-prototypes, SABC, KL-FCM-GM and IWKM algorithms by 14.4%, 38.3%, 9.7% and 9.2%, respectively (see Table 4). Clustering accuracy *r* was 0.779 for the heart disease dataset using the MD-MWFCM (see Table 5). This was 23.3%, 23.4%, 12.6%, and 12.6% more clustering accuracy than that of the k-prototypes, SABC, KL-FCM-GM, and IWKM algorithms. The comparison of the maximum accuracy achieved by five clustering algorithms for the Iris data set is as shown in Table 1:

The comparison of the maximum accuracy achieved by five clustering algorithms for the Soybean data set is shown in Table 3.

The comparison of the maximum accuracy achieved by five clustering algorithms for the Heart Disease data set is given in Table 4

A comparison of the common algorithm for clustering mixed categorical and numeric data set is shown in Table 5

The k-prototypes algorithm has its advantage in a global optimal search, but falls short in losing information about multiple memberships of a data point (refer Table 5). The SBAC algorithm uses the same agglomerative algorithm to construct a dendrogram to display the similarity (or

TABLE 3: Comparison of the maximum accuracy achieved by five clustering algorithms for the Soybean data set.

Algorithm	Clustering accuracy (r)
k-prototypes	0.856
SBAC	0.617
KL-FCM-GM	0.903 ($\alpha = 1.8$)
IWKM	0.908
MD-MWFCM	1.000 ($\alpha \in [1.1, 1.2]$)

TABLE 4: Comparison of the maximum accuracy achieved by five clustering algorithms for the heart disease data set.

Algorithm	Clustering accuracy (r)
k-prototypes	0.546
SBAC	0.545
KL-FCM-GM	0.653 ($\alpha = 1.3$)
IWKM	0.653
MD-MWFCM	0.779 ($\alpha = 2.2$)

difference) level of a pair of data points with numeric or nominal attributes. Its limitation is to give larger weight to uncommon attributes which may lead to its unreasonable significance in clustering. The KL-FCM-GM algorithm with Gath-Geva theory is more suitable for clustering numeric data with Gath distribution. Although the IWKM algorithm performs relatively better than k-prototypes, SBAC and KL-FCM-GM algorithms in mixed numeric and categorical data types, it is over sensitive to cluster center initialization for different attributes. The proposed MD-MWFCM clustering algorithm has shown its higher performance than these clustering algorithms for mixed data. It has introduced a different representation of cluster centers for different data types, i.e., numeric data center for numeric data type and categorical data center for the categorical data type. For numeric and categorical data in biomedicine, techniques for dealing with categorical data are usually more difficult than those for numeric data. The k-modes, as a classical categorical clustering method has demonstrated its advantages. Compared with k-modes, the MD-MWFCM clustering algorithm has been improved in following aspects. First is that fuzzy membership has be applied in MD-MWFCM clustering algorithm to reflect the relation degree of a data point to each cluster center while k-modes only considers hard fuzzy partition. Moreover, fuzzy membership also has been both combined into the dissimilarity measurement and objective function process with weighting. The MD-MWFCM clustering algorithm uses different dissimilarity

Table 5: A comparison of the common algorithm for clustering mixed categorical and numeric data set.

Clustering method	Analysis	Limitation
k-prototypes [10]	Step 1. Select several data points as the initial cluster centres randomly Step 2. Allocate the rest data points with the highest similarity to the relevant cluster centre Step 3. Update the cluster centres after each data allocation Step 4. Recalculate the dissimilarity of each data point to each cluster centre Step 5. Repeat steps 2 to 4 until reaching the optimal cluster allocation	Data can only be allocated into one cluster, losing information about multiple membership; Sensitive to initialization of cluster centres; No discretion of varying weights of an attribute in different clusters
SBAC [11]	Step 1. Create a cluster that contains a pair of data points with the highest similarity Step 2. Add another data point and compare its similarity with the initial two data points. If the similarity is higher than that of the two original ones, add to the cluster; otherwise treat the data as a new cluster centre Step 3. Repeat Step 2 until all data points have been located to a cluster	Uses the same method to measure both numerical and categorical data; Sensitive to initialization of cluster centres; No discretion of varying weight of an attribute in different clusters
KL-FCM-GM [13]	Step 1. Use gath-geva theory to allocate fuzzy membership of all data points Step 2. Update weight of each data point in the allocated clusters Step 3. Retest the distance of data point to cluster centre Step 4. Estimate the cluster distribution parameters of gath-geva formula again Step 5. Repeat steps 1 to 4 until the objective function has reached the stop condition	Sensitive to initialization of fuzzy membership
IWKM [14]	Step 1. Select several data points as the initial cluster centres randomly and assign weight to each Step 2. Compute fuzzy membership matrix based on the values of the initial cluster centres and their weights obtained in Step 1 Step 3. Update cluster centres with the values of fuzzy membership matrix and weights in Step 2 Step 4. Update the weights with the values of fuzzy membership matrix and cluster centres in Step 3 Step 5. Repeat steps 2 to 4 until the objective function has reached the stop condition	Sensitive to initialization of cluster centres

measure individually for numeric and categorical data in a cluster. This can fully use fuzzy membership in clustering vague data across the boundaries of clusters. The objective function calculates dissimilarity by data types and weights of attributes in the iteration of clustering. In the partitioned clustering algorithm, the initial position of the cluster center has a significant impact on the clustering results. The randomness of this initialization method is too strong, which leads to great fluctuation of clustering effect, and the same clustering results are difficult to reproduce. The improved cluster center initialization method in MD-MWFCM clustering algorithm is proposed. The initial centers from the results in the minimum threshold method can help avoiding the randomness problem of initial cluster centers in a random way. The minimum threshold method can fix the cluster centers in a possible defined range. All these have enabled the MD-MWFCM to outperform the other four algorithms in clustering all three datasets, only numerical, only categorical, and mixed numerical and categorical datasets.

(1) The Optimal Fuzzy Coefficient α Value in Accordance with the Maximum Clustering Accuracy **r**. When $\alpha \in [1.0, 3.0]$, the Iris dataset achieved its optimal clustering accuracy of 0.967 as shown in Figure 2).

When $\alpha \in [1.0, 10.0]$, the Soybean dataset achieved its optimal clustering accuracy of 1.000 as shown in Figure 3

When $\alpha \in [1.0, 10.0]$, the heart disease dataset achieved its optimal clustering accuracy of 0.779 as shown in Figure 4

Because overlap degree between the clusters in a data set is mainly impacted by the fuzzy coefficient α, the degree of overlap reflects the distribution of data points in each cluster which can be indicated by calculating clustering accuracy r. Therefore, the clustering accuracy r in the MD-MWFCM algorithm is mainly impacted by α (see Figures 2–4) instead of the numerical value fluctuation in the Iris dataset or the categorical values in the Soybean dataset. It has been an ongoing challenge to select the value of α in fuzzy data analysis. Several researchers have attempted to address this challenge. These attempts all use mathematical programming in selecting the values of a fuzzy correlation coefficient, and it is not easy to propose further applications for researchers who don't have a strong mathematical background. The authors suggested that the optimal clustering accuracy r is achieved when α is within 10.0, especially when $\alpha \in [1.5, 2.5]$ [14]. Our study has contributed empirical evidence to support their suggestion. By starting the experiment with the αvalue of 1.0, and incrementing it at 0.1 interval, we can accurately pinpoint the value α with
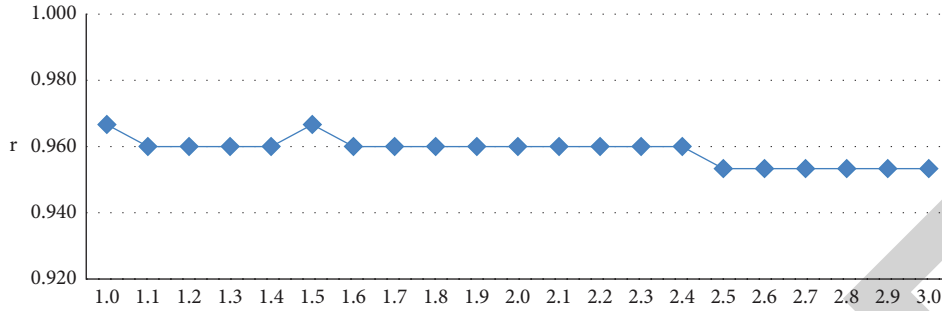
Figure 2: The optimal accuracy (**r**) of the MD-MWFCM with varying fuzzy coefficient α for the Iris dataset.
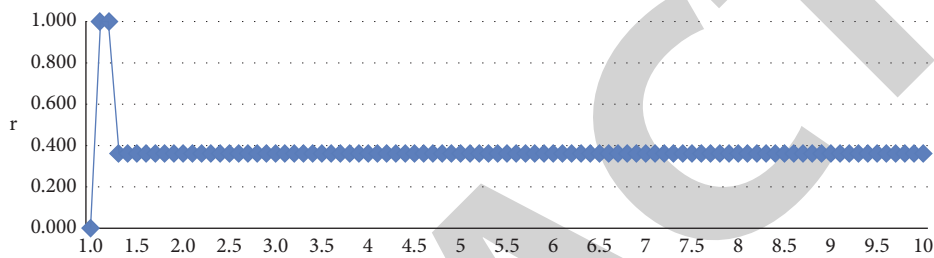
Figure 3: The optimal accuracy (**r**) of the MD-MWFCM with varying fuzzy coefficient α for the Soybean dataset.
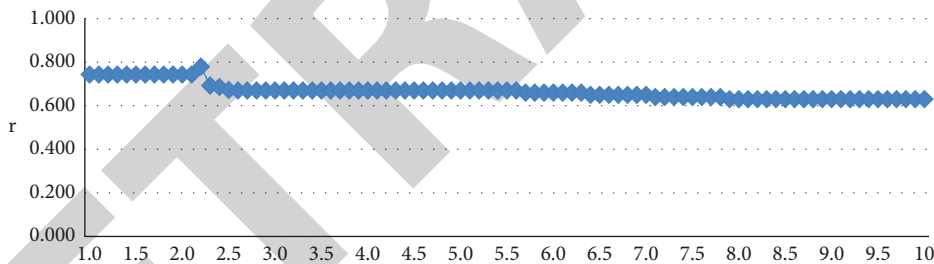
Figure 4: The optimal accuracy (**r**) of the MD-MWFCM with varying fuzzy coefficient α for the heart disease dataset.

accuracy to 0.1 for the corresponding optimal clustering accuracy r. Every dataset has its own optimal fuzzy coefficient value suggested by the corresponding optimal clustering accuracy r (see Figures 2–4). Before achieving the aim of the objective function, the clustering accuracy r is not substantially impacted by the changing value of fuzzy coefficient α; this suggests that the performance of the MD-MWFCM algorithm is reasonably stable. According to article [3], one performance criteria of a clustering algorithm is its convergence speed of the clustering algorithm. According to article [4], being able to achieve the optimal clustering accuracy with $\alpha < 10.0$, is an indication that the objective function can converge at reasonable speed. The convergence was achieved when $\alpha = 1.5$ for the Iris dataset (Figure 2), $\alpha \in [1.1, 1.2]$ for the Soybean dataset (Figure 3) and $\alpha = 2.2$ for the Heart Disease dataset (Figure 4). Because the objective function achieved convergence with small fuzzy coefficient values $(\alpha < 3.0)$ in all three data sets, it suggests that the MD-MWFCM algorithm has high performance speed.

4. Conclusion

To address the limitations of FCM, this study proposes a novel multiple weighted fuzzy C-means for mixed data (MD-MWFCM) clustering algorithm. The MD-MWFCM algorithm presents a novel clustering FCM algorithm by treating numeric and categorical data individually with their respective cluster center representation, dissimilarity measurement and objective function aspects. For cluster center representation, mean is used for numeric data and fuzzy center is used for categorical data. Based on the different cluster center representations, it uses different dissimilarity measurement methods to compute the distance of a data point to various cluster centers. For dissimilarity, numeric data are calculated by the distance and weight of a data point to each cluster centre, categorical data are measured by the frequency of occurrence and weight of a data point to a cluster. The objective function is computed by the dissimilarity measurement. Moreover, the MD-MWFCM

Research Article

Analysis of Performance Improvement of Real-time Internet of Things Application Data Processing in the Movie Industry Platform

Yang Meng [1,2]

[1]*Macau University of Science and Technology, Faculty of Humanities and Arts, 999078 Macau, China*
[2]*Communication University of Shanxi, Shanxi 030619, China*

Correspondence should be addressed to Yang Meng; 1909853nam30001@student.must.edu.mo

The goal of this study is to plan and develop complete strategies to improve the performance of film industry. The primary objectives of this study are to investigate a dataset generated by a IoT application and the nature of the data forms obtained, the speed of the data arriving rate, and the required query response time and to list the issues that the current film industry faces when attempting to handle IoT applications in real time. Finally, in film industry platforms, high performance with varied stream circulation levels of real-time IoT application information was realized. In this study, we proposed three alternative methods on top of the Storm platform, nicknamed Re-Storm, to improve the performance of IoT application data. Three different proposed strategies are (1) data stream graph optimization framework, (2) energy-efficient self-scheduling strategy, and (3) real-time data stream computing with memory DVFS. The work proposed a methodology for dealing with heterogeneous traffic-aware incoming rate of data streams Re-Storm at multiple traffic points, resulting in a short response time and great energy efficiency. It is divided into three parts, the first of which is a scientific model for fast response time and great energy efficiency. The distribution of resources is then considered using DVFS approaches, and successful optimum association methods are shown. Third is self-allocation of worker nodes towards optimizing DSG using hot swapping and making the span minimization technique. Furthermore, the testing findings suggest that Re-Storm outperforms Storm by 20–30% for real-time streaming data of IoT applications. This research focuses on high energy efficiency, short reaction time, and managing data stream traffic arrival rate. A model for a specific phase of data coming via IoT and real-time computing devices was built on top of the Storm platform. There is no need to change any software approach or hardware component in this design, but only merely add an energy-efficient and traffic-aware algorithm. The design and development of this algorithm take into account all of the needs of the data produced by IoT applications. It is an open-source platform with less prerequisites for addressing a more sophisticated big data challenge.

1. Introduction

Digital technology is altering scientific practice. Digital imagery, sensors, analytical apparatus, and other techniques are becoming increasingly important in many areas of science for experimental and observational study. Usually big data is one of the technologies to deal a massive number of datasets [1, 2]. Currently, two different processing/ computing platforms are there to deal big data. One is a batch computing platform like Hadoop. The second one is stream computing like Storm. Because it takes data in a

stream form from collections of hardware or software sensor and computes continuous time data streams, stream processing is a key feature of this program on real-time online high-speed streaming data. For processing enormous volumes of streaming data, real-time stream computing is required. The future consignment is based on common task structure experiential in several Internet of things (IoT) fields for decision-making in real time, and the input data streams are obtained from real IoT applications observations from various applications like smart hospitals, smart cities, and entertainment like film industry. The collaboration of

this application is full by the integration of business, consumer, and business and industry Internet access, and industrial IoT consumer interactions is used by IoT applications [3–5].

In big data systems, the preparation of high-speed continuous information in stream processing required for processing is also a noteworthy goal. In the fault-tolerant population-stream registration programming mentioned by Neumeyer and Robbins(2010), which mimics S4, the shortcomings of group processing are addressed by equal-flow computation with low inertia and acceleration, but the full-population-stream processing of both platforms has drawbacks. By conducting a thorough examination of existing rules, it was discovered that the present film and television engines are not meeting the demands of IoT data streams. Unknown energy consumption, reaction time period, and traffic velocity data entering the streaming data engine are provided by the IoT data element. IoT is expanding and making immense measures of constant information; it is a major testing assignment in IT industry. Stream registering is reasonably the quickest and most proficient answer for getting profitable data from enormous information. In addition, numerous information streams from divergent information sources may shape a mix of the information sorts that might be inconsistent [6–8].

The goal of this study is to plan and develop complete strategies to improve the performance of film industry. The environment for IoT application datasets makes a platform towards using data stream optimization, energy-efficient data stream dynamic scheduling, and memory DVFS approaches to reduce energy consumption. Future IoT advancements will handle highly dispersed IoT applications that necessitate a high level of distribution and process at the network's edge by using platform that would provide compute, storing, and data networks between edge devices or computing data centers. These systems will enable emerging Internet of Things applications that require real-time latency. This study effort begins by outlining the mathematical relationship between energy use, reaction time, and overall asset use in the film business. It finds out how to satisfy the reduced response time and the high energy efficacy targets, displaying DSG in film industry conditions by alluding to the appropriated stream figuring hypotheses, distinguishing that the vertices weight are vertices in DSG, and achievement energy utilization of a distribution conspire for a DSG at a particular information stream speed, and doling out assignments by the energy proficient heuristic movement mindful planning strategies.

The study's goal is to increase the performance of real-time IoT data processing on a film industry framework. The main goals of this research are to evaluate a dataset generated by a popular IoT application and the nature of the information obtained, the velocity of responsive service rate to computing, the query-processing time needed, and to list the issues that the current film industry system faces when dealing with IoT applications in real-world time. The high performance of real-time iot application information with different flow circulation degree is realized on the film industry platform.

2. Methods

2.1. Data Stream Graph Optimization Framework.
Approaches to constructing a BDSC platform based on the data streams graph optimization method make it possible to optimize DSG via critical route removal and parallelism. Optimizing an application's scheduling strategy DSG to assets is being considered, but how to progress the DSG is being disregarded. To attain high consistency, it is critical to first obtain a good image of the altered state of the DAG and then select which vertex of a DAG should indeed be rescheduled. Before submitting the graph to the Storm stage, the client plans the structure of the DSG based on the capability of the application. To ensure high stability, establish a fair picture of a DAG's changing state before deciding which vertex of the DAG must be rescheduled. More crucially, understanding how to boost framework strength while assuring makes span reduction and controlling superior, and response time trade off in a productive and practicable method is required, which is absent in the bulk of current scientists in BDSC situations [9–12].

2.1.1. System Architecture.
The four stages of the Storm computing platform are responsible for real-time data streaming (Figure 1) which are task assessment, scheduling phase, processing phase, and storage area. It demonstrates DSG optimization utilizing two unique strategies: critical route elimination, which avoids the crucial route to modify the latency of the produced data stream, and data stream parallel processing, which approaches heavy nodes of the are computing data stream operating in parallel (Figure 1).

2.1.2. Experimental Set-Up.
A model resolve is the process of creating a simulation environment for real-time computing. To create a simulation environment, hardware requirements are utilized. To test the efficacy of the suggested paradigm, an experimental setting with high-speed network access was created. Intel i7 CPU, 16 GB RAM, 1 Gbps network access, and 104 core workstations were used to evaluate the suggested concept. Two proposed methodologies were applied to tuples that comprised a stage of the Storm framework. Ubuntu server 14.01, Storm 0.10.0, Java 1.8.25, Zookeeper 3.4.0, and Python 3.0 are the software requirements for computing the results. A cluster's performance is evaluated using real-world data tests. The cluster comprises of 18 computers, one of which is assigned as a main node that runs Storm Nimbus, one as a Zookeeper node, and the remaining 16 as subnodes. Each computer runs with Linux Ubuntu Server 13.04 and is equipped with twin 8-core Intel Core (TM) i7-4790 processors running at 3.6 GHz, 16 GB of memory, and 1 Gbps NIC (Figure 2).

2.1.3. Optimizing.
Before the graphs have been submitted to the phase on the Storm platform, the client defines the structure of the DSG based on the capacities of the application [13]. DSG has two variables such as $G = (VG, EG)$; in this group, the vertex is treated as DAG, which is a
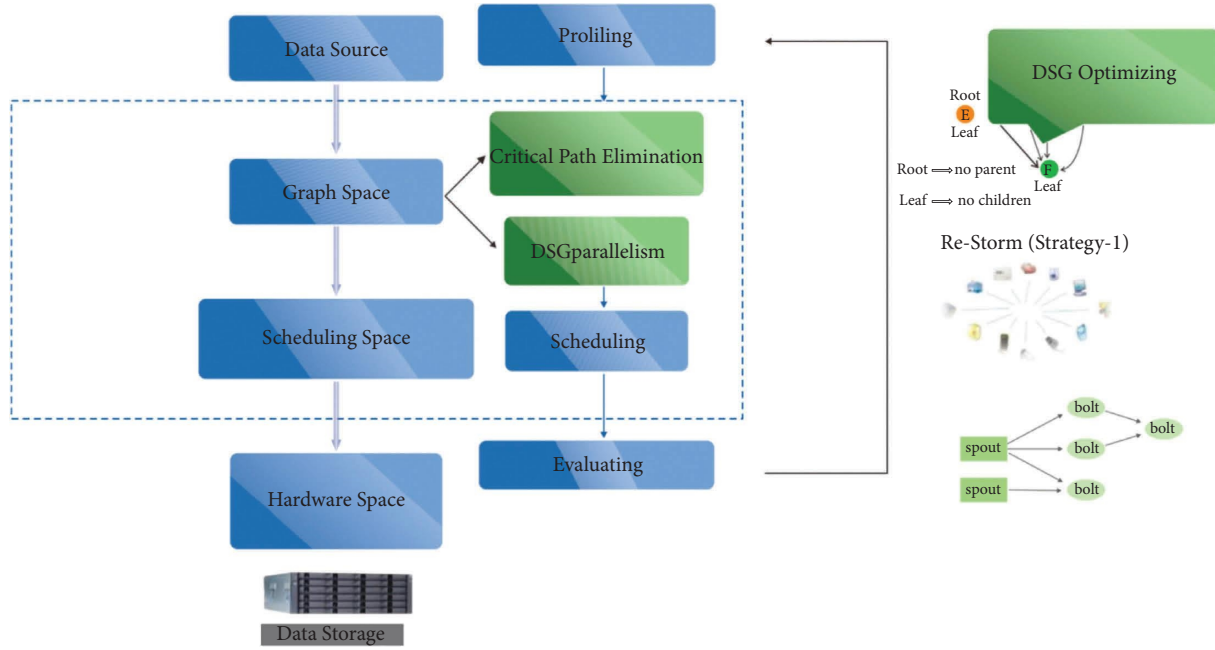
FIGURE 1: Computing data stream operating in parallel.

Directed   Instant Graph   Logical Graph
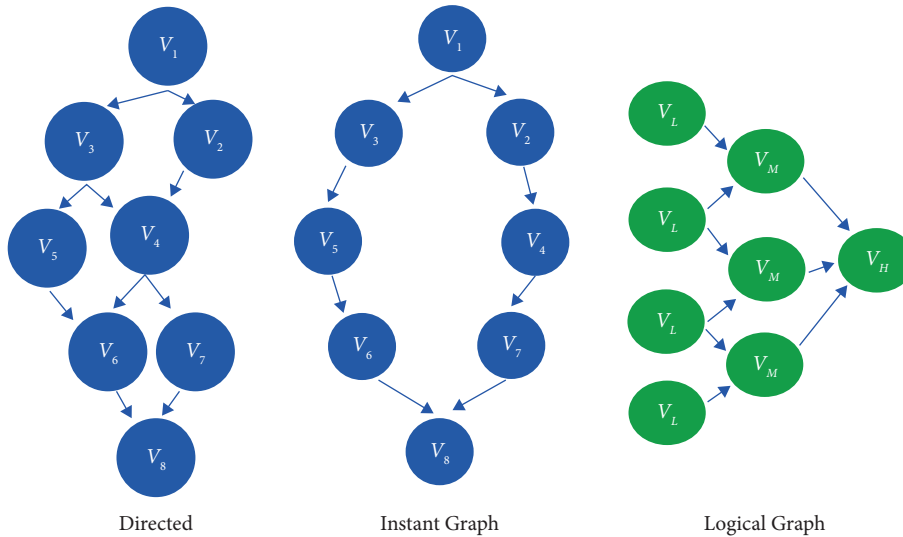
FIGURE 2: Model resolve.

consequence by the DAG. Each G continues using two parameters G = (VG, EG), where VG is a group vertex and EG is a group edge, and subgraph Gs, $\forall$ Vi ϵ VGs than $\forall$ VG ε VGs, the way it is travelling the DAG route. If S $\neq$ E, then the starting and ending points of a (Vs, Ve) are the same, the graphs are not directed, and it usually indicates a null vertex/ node. Topological sort (TS) is also another type of graph that does not have cycle structures. DAG stands for topology sort order. The DAG considers partitioning a graph based on TS by separating the vertex of a graph. A partition graph (GP) is a topology-based partitioning vertex graph GP = {GP1, PP2, PP3 . . . . .}. For each partition, GP1 = {v1, v2, v3, . . . } $\in$ G and GP2 = {v1, v2, v3, . . . } $\in$ G. It is a subgroup containing $\forall$ $i \neq j$ i, $j \in$ (1, n), then GPi $\cup$ GPj = $\varnothing$ $\cup$ GPi $\neq$ n 1 VG and G = (VG,

EG). Each vertex and edge contains some tuples VG = (idv, fv, cvi, v, ov) and EG = (ide, ce) P(Vs, Ve), where graph vertices = {V1, V2, V3, ... Vi}, graph edges = {E1, E2, E3, ... Ej}, start vertex and end vertex = Vs, Ve identification of vertices, function, computing cost, data input streams, output data stream idv, fv, cv, iv, ov 52 ide, Ce = directed edge identification, directed edge communication cost. In the graph shown in Figure 2 , it starts vertex from V1 and ends at V8, and it is not encompassing circles and it ϵG {V1, V2, V3, V4, V5, V6, V7, V8}, and in this one subgraph, we assume, for example, {V1, V3, V5, V6, V8} $\in$ G, and there are two routes in this above graph {V1, V2, V4, V7, V8} and {V1, V3, V5, V6, V8} $\in$ G. Furthermore, TS is depicted in the graph (Figure 2).

2.2. Energy-Efficient Self-Scheduling Strategy (EESSS)

2.2.1. Framework. In Storm processing area, the model for observing digital streaming graph, optimization model for DAG, adaptive streaming data deviation model, energy-aware scheduling of DAG approach, and traffic-aware resource rescheduling method for the DAG model work to care for constant traffic rate for scheduling intends on top of the open-source Storm platform. The storm is indeed a BDSC system that is distributed. To disperse data streams between several instances of a vertex, the energy-efficient traffic awareness resource scheduling paradigm is employed. The DAG-based energy-aware scheduling approach is utilized to give a DAG self-scheduling strategy based on traffic rate that is created by a user and is in a traffic rate stream. The DAG model's traffic awareness based resource rescheduling is used to reschedule a DAG in an energy-efficient manner and to achieve high energy efficiency in a constant fashion [14–18] (see Figure 3).

2.2.2. System Architecture. The flow of system is considered either by user or hardware generation of data streams as well as its formation as a graphical form, and then it is transmitted by the Storm computation, and this process using predefined task scheduling by round Robin made reference from Zong et al. for enhancing the performance of standard scheduling strategic approach which is amended as an energy-efficient traffic-aware resource allocation method (as shown in Figure 4) [18].

2.2.3. Experimental Set-Up. By changing the default scheduling method in Storm's IScheduler surface and gaining correct results, efficient energy self-scheduling methods are used to improve performance (by optimizing energy efficacy and reducing reaction time through varied levels of traffic of the data streams). The Storm platform simulation environment is built in fully functioning parallelism, fault tolerance, and distribution of the latest version of software, i.e., Storm 0.10.0. A 4 core Intel 13 processor 2.00 GHz 64 Bit CPU, 16 GB memory, and 512 Mbps network access are required for virtual machines. 4 core two PCs, each with a 10 TB external storage capacity, are joined together. A Linux server is installed on each system (Ubuntu version 14.01). The software components listed below are often built and used in combination with Java 1.8, Zookeeper 3.4.0, and Python 3.0. Furthermore, on the Storm platform, all upgraded scheduling techniques replace the default scheduling approach with improved and efficient traffic-aware scheduling for energy. In Storm UI, the output is being watched. For performance purposes, the average tuples computation is employed. This used Storm's default timing technique to measure the process time of every tuple. Storm UI can gather such information, but then it only displays the 10 minute average. According to Zhao Zhibin et al., the recommended approach trial and execution takes about one minute rather than averages, providing us with far more precision in real-time performance estimate (2008). During the investigation, Ubuntu Linux used the NTP protocols standard to synchronize worker nodes [19].

2.3. Real-Time Data Stream Computing with Memory DVFS. DVFS is a popular technology to scale voltage frequency according to the application precedence at the CPU level. In this work, this technique is apllied to the memory level to reduce energy consumption and improve performance.

2.3.1. Experimental Set-Up. Adding frequency scaling-based control algorithm for improving energy efficiency is done. It minimizes the application's energy usage while increasing its efficiency. The frequency-based control mechanism is software-based. Figure 5 depicts the data stream control method. It is to handle the large data stream computing environment while processing data from IoT applications. It experiences difficulties; thus, it is searching for more gap filling to assist it in overcoming these difficulties. The Storm framework in the stream computing is open source and is developed to address the most pressing demands of the current streaming data component. The scheduling mechanism used in this one is Round Robin by default. It is undesirable as data transmission is rather high at the slow point, and energy usage is quite significant for that reason. Data were generated by devices (Figure 5).

2.3.2. Optimization. Sun et al. used the DVFS approach to the enormous data leaking group and scientific proof presented with each on/off chips doling out workloads [20]. The load of an errand is defined as the total of the CPIs of all bearings further towards the path stream of the venture. A variety of component factors influence the task load, including the on-chip halt cycle number owing to data reliance or the branched miss forecast, and the off-chip log jam phase checks due to I/D TLB miss or I/D store miss. The CPU waits until the requested memory exchange is accomplished during an off-chip access. As a result, the processor clock cycle during an off-chip is altered. To comprehend the load rot framework, a few definitions are required.

Definition 1. Won is the number of clock cycles necessary for the CPU to complete the set of on-chip instructions, which are made entirely of the CPU. The performance time required to complete Won, T on, fluctuates according to CPU frequency, f cpu, and is calculated as T on = Won/f cpu.

Definition 2. Woff is the number of external clock cycles required to conduct the set of off-chip accesses. It is worth noting that the CPU delays until the peripheral memory operations are completed. The execution time necessary to complete Woff, T off, is computed as a function of the outward memory clock cycle frequency, f ext as T off = Woff/ft = −logf (T o f 2 Wo f 2) ex, $x ≠ 0$ and −Wo f 2 ≠ 0 ∅, $x = 0$ and −Wo f 2 ≠ 0 and −logf (T o f 2 Wo f 2) ≠ 0 $t ∈ ℝ$, $x = 0$ and −logf (T o f 2 Wo f 2) = 0 and −Wo f 2 ≠ 0. Based on defined in equations 5.1 and 5.2, Won and Woff is written as Won = N.CPUIavg on and Woff = M. CPI off avg (5.1),
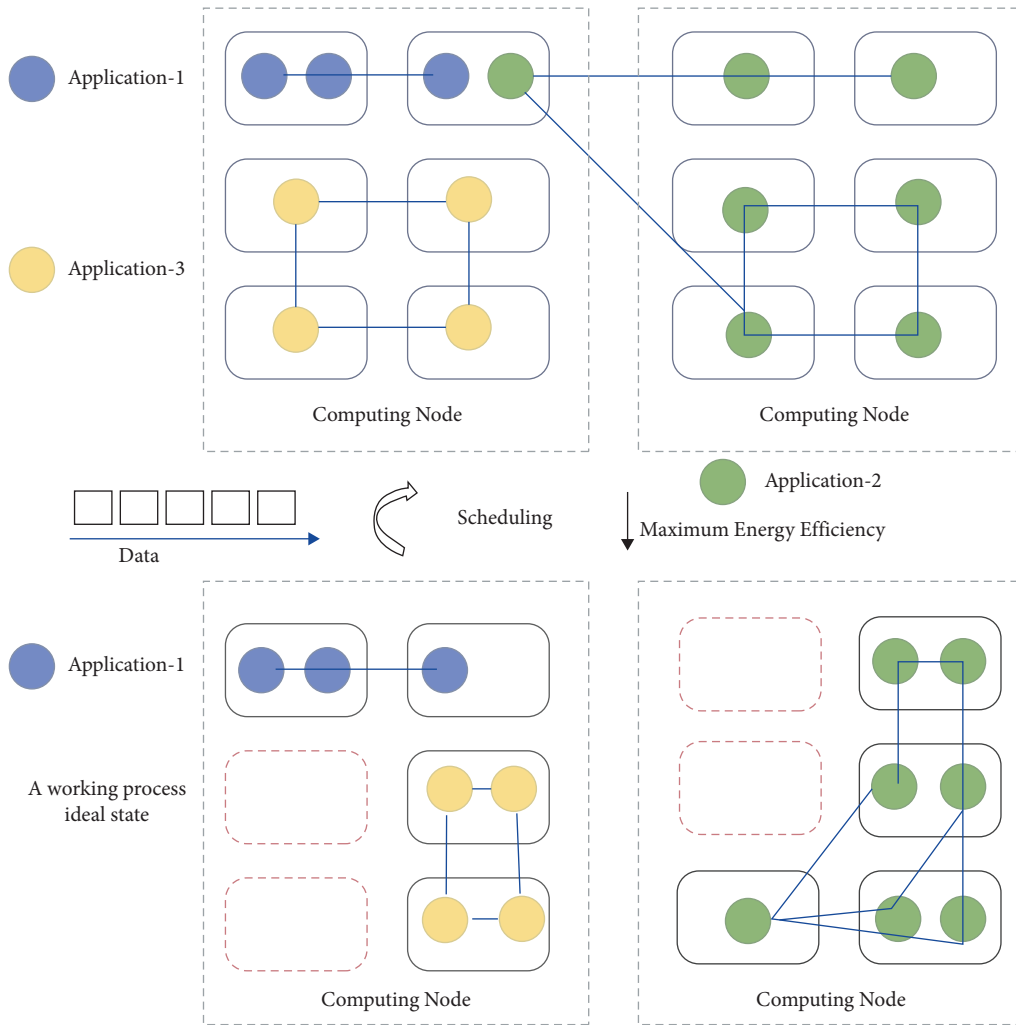
FIGURE 3: Framework.
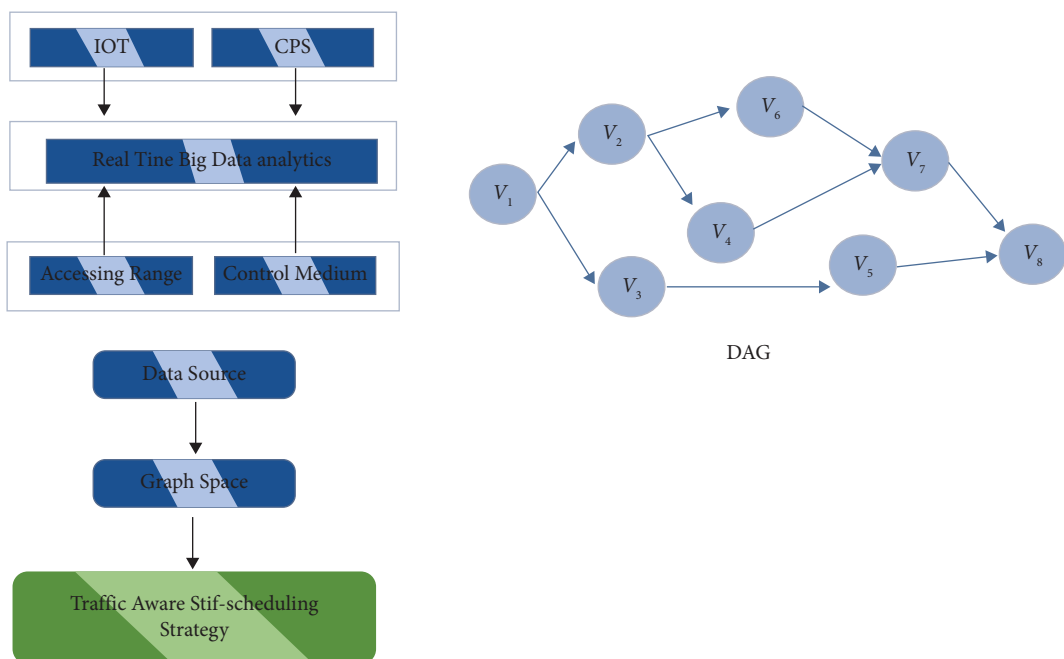
FIGURE 4: System architecture.

Input: Real- Time IoT streaming data

Output: Optimized Solution with frequency based task allocation

Srep 1: Bandwidth required data stream

Step 2: Bandwidth available to process task

Step 3: Average Bandwidth thread assigning task

Step 4: if Bandwidth < Tf_b then

Step 5: set memory frequency to f_b

Step 6: else if T$f_b \leq$ Bandwidth < Tf_c then set memory frequency to f_2

Step 7: else if T$f_c \leq$ Bandwidth then

Step 8: set memory frequency to maximum frequency f_a

Step 9: end if

Step 10: end while

FIGURE 5: System architecture.

where CPI avg on represents the average number of CPU clock pulses per on-chip instruction, M represents the number of off-chip access, and CPI avg off represents the average of external clock cycles per off-chip access. The execution time, T, for a job is determined using these two definitions as $T = T$ on $+ T$ off $= N$.CPUIavg on fcpu $+ M$. CPI off avg fext (5.2) When the CPU frequency changes, the change in T is solely due to T on $(\Delta T \Delta f$ CPU$) = \Delta T$ on Δf cpu, ΔT off Δf cpu 0 (5.3) Lemma 1: Each degree of energy consumption is based on computing each job and reaction time based on a scheduling element, and two tasks A and B are considered. Proof: Assume 89 Accept to make a change in which you have two distinct ideal schedules A and B. Consider a third schedule U wherever for all the tasks i, xi(U) = (xi(A) + xi(B))/2. Now privilege that F(U) ≤ (F(A) + F(B))/2 = F(A) = F(B) and E(U) < (E(A) + E(B))/2. As a result, neither S nor T are optimal for one may enhance the plan by investing A − E(U) energy into occupation n in U to show evidence of better reaction time than F(U). This opposes the model of S. To demonstrate F(U) ≤ (F(A) + F(B))/2 = F(A) = F(B), consider a certain task b. Then there is some work "a" Cy (U) = rx + ∑ (U) yi = a. So by the description of U, Cy(U) = rx + ∑(xi(A) + xi(B))/2y $i = x$. But in S, let it necessarily be the case that Cb (A) ≥ ra + ∑ (xi (A) yi = x (5.4). For A, must procedure jobs x finished y among response time rx and Cy(A). Similarly, Cy (B) ≥ rx + ∑ xi(B) y $i = x$ (5.5). By taking the average of these two equations, (Cy(A) + Cy(B))/2 ≥ ra + ∑ (xi (A) + xi(B))/2 y$i = x$. (5.6). The right-hand side of this inequality is precisely Cy (U) in this processing. Hence, (Cy(A) + Cy(B))/2 ≥ Cy(U). Because y was picked at random, it follows that F(U) ≤ (F(A) + F(B))/2. Note that the function f(x) = 1 x $\alpha-1$ is a rounded function when $\alpha > 1$, and $f (x + y$ 2$) < ($f$(x) + f$(y))/2. It thus immediately follows E(U).

3. Results and Discussions

All of the findings are obtained from various sources in order to verify the working nature of the offered technique. Streams submitting at various traffic levels are on top of it, testing to see if it is appropriate for all situations and comparing the findings to different qualities. The principal input type is data generated by IoT devices, which is perfect for sampling real-time and high-velocity data but also a difficult task in today's IT industry. The proposed method is for efficiently providing greater control over IoT-related data processing via big data platforms which meets this aim, meaning that it is suitable for readily computing various types of high-velocity data.

In this paper, we presented a new application standard to measure BDSC in the IoT environment. Data phases, like stream processing system (SPS), are required for IoT applications' high speed control needs, and recommended work overload calculates their viability using basic tasks found in IoT applications, as well as entirely practical implications for the fact-based outline and predictive investigative process. These are combined with two trustworthy information sources from the urban IoT testing and transportation businesses. The proposed standard for the widely used Apache Storm SPS, as well as the implementation steps, has been accepted. A task scheduling planning calculation for managing massive data streams in mobile Internet service is provided to establish parallel machine execution, and the streaming query graph is functioned to determine each edge weight. The remodeling findings show that using the appropriate number of logic machines reduces the response time of framework substantially, and scheduling several tuples at once reduces framework connection switching. The calculating approach used in this study can increase the productivity of massive data stream processing in portable one. The suggested data stream optimization has indeed been accepted as the benchmark for the widely used Apache Storm SPS and the execution methods introduced. A dynamic programming planning calculation for big data stream handling in mobile Data Internet access is offered to create parallel machine execution, and the streams query graph is worked to determine each edge weight. The remodeling findings show that using the appropriate number of logic machines reduces framework response time substantially, and scheduling several tuples in one go minimizes framework connection switching. This study's calculation has the potential to

increase the efficiency of enormous data stream handling in mobile Internet access. Reduced scheduling rates, on the other hand, will lead to IoT implementation [21–25].

3.1. Energy-Efficient Self-Scheduling Strategy (EESSS).

All of the findings are collected from various sources in order to verify the effectiveness of the suggested approach, with streams submitting at various traffic levels on top of it, testing to see if it is appropriate for all situations and comparing the findings to different qualities. The principal input type is IoT-produced data, which is appropriate for sampling real-time and high-velocity data but also a difficult task in today's IT industry. The proposed method is for efficiently providing improved control for Internet-of-things data computing by using big data frameworks that meets this aim, meaning that it is suitable for readily computing various types of high-velocity data. Obtaining data samples of real-time IoT creation from Shahrivari's CityPulse Database Collection (2013). There are several types of datasets accessible, such as pollution, weather, and road traffic which developed a virtual environment for controlling data speed with various traffic mediums.

When compared to the present model, evaluating the suggested model with appropriate varied traffic mediums would yield better performance. First, allocating a low traffic medium, then making numerous alterations and eventually achieving peaceful outcomes in all aspects. Those graphs are provided below. We improved energy economy and reduced reaction time by varying the traffic volumes of information streams. The graphs show the outcomes of how to use energy at the system level while lower traffic levels are evaluated. Figure 4, 6 depicts a 0–250 tuple which constructed a variable of tuples on submitting somewhere in the middle of the range at the responding time and energy consumed with both platforms under the identical circumstances. Re-Storm outperforms both Storm and Re-Storm in all scenario test situations, demonstrating that it is well suited for IoT-sensing data. Online ongoing information is indistinguishable, divergent arriving rate, and it does not have a consistent activity detecting medium. It was examined what all significance are there for upgrading stream diagrams and energy proficiency requirements for BDSC condition. The proposed approach considered two variables for enhancing their execution proficiency. Initially, it is adjusting their planning system with unessential activity stream support, and second is enhancing their diagrams utilizing basic way disposal to keep up a voting demographic for various movement medium information, besides, updating 20–30% proficiency in stream figuring. At long last, it makes a colossal effect on general all BDSC condition obtaining exceptional performance throughout the whole big data platform, with a focus on real-time as well as IoT information. This research concentrates on improving energy efficiency, fast reaction time, and controlling the arrival rate of data stream traffic. On top of the Storm system, a model for a specific phase of data arriving from IoT and real-time computing was built. This design does not require any changes to the software or hardware; simply add an energy-efficient and traffic-aware algorithm. The design and development of this algorithm take into account all of the needs of the data produced by IoT applications. It has less prerequisites for dealing with a more complicated big data challenge and is an open-source system [26, 27].

3.2. Real-Time Data Stream Computing with Memory DVFS.

This approach presented the essential tradeoffs in memory recurrence scaling and played out an underlying assessment utilizing a straightforward and natural calculation. However, more work stays to be finished. To begin with, a basic system and a vast plan space, both plays a role in measuring and anticipating the effect on execution and on expecting the future effect of memory recurrence changes. Additionally, work can explore both the estimation and forecast parts of this issue and portray how different sorts of workloads react to expanded memory inertness. This paper also examined the interaction between memory storage scaling and CPU voltage or frequency scaling (DVFS). Positively, the two devices could exchange signals. It is also possible that higher productivity gains are possible under combined control as when the two work independently. At long last, it is thought to be just SPEC CPU2006 in this work; assist assessments are important to evaluate execution affect in different workloads. In this paper, a model was suggested to analyze memory frequency/voltage scaling in order to maximize energy efficiency and minimize memory power. We describe a control technique that decreases memory frequency while reducing performance effect, based on the fact that a large amount of memory system power is frequency dependent. The essential discovery is that changing memory frequency has no effect on memory access latency when memory bandwidth usage is low. By monitoring memory bandwidth consumption, the suggested control method raises memory frequency when utilization exceeds a certain threshold, hence limiting the performance effect. In this way, DVFS can be a useful energy-saving solution, especially when memory bandwidth consumption is modest [28].

4. Conclusion, Future Work, and Significance of the Contribution

4.1. A Data Stream Graph Optimization Framework.

It maintains a consistent high efficiency and a short response time. It decreases the utilization no. of computing nodes. The quantification of distribution scheme for a DSG is energy efficient at a given information stream velocity and merging of heavily weighted vertices on large DAG to perform logical splitting to maximize the energy efficiency, altering their scheduling strategy with inappropriate traffic streams and optimizing DSG using vertices weight-based exclusion. Moreover, the upgrading approach against the existing platform enhanced efficiency shown by 20–30%. Finally, it makes a huge effect on the whole BDSC environment.

4.2. An Energy-Efficient Self-Scheduling Strategy.

This technique presents Re-Storm, a revised Storm streaming engine, as well as an energy-aware stream congestion consolidation

mechanism for allocation of resources. As per the arriving size of the data stream weight of each vertex, it should allocate resources as per online data stream traffic rate. Finally, achieving low response time and high energy efficiency with dissimilar traffic levels of streams in BDSC is called as Re-Storm.

4.3. Real-Time BDSC Platforms towards Improving Efficiency. A new main memory-based control algorithm strategy is proposed for observing memory bandwidth based on task strength adjusting, utilizing its frequency/voltage scaling to minimize performance impact. The power consumption of memory is a significant module of system power. Reduce power consumption memory level as it has high effect and impact on the overall memory level computation system performance. About19% of average gain is achieved against the existing strategies in the evaluation system. Scaling memory voltage/frequency can reduce power utilized by memory with a minimal system performance effect which yields average system energy with a reduction by 2.4%, achieving 0.4% of average memory power reduction. These three proposed strategies are added into the existing BDSC platform towards improving the efficiency of real-time computing of the IoT Applications.

5. Additional Points

The limitations of study is that assignments on web-based advancing and engineering which necessitates for the booking component by reallocating the basic vertices continue the basic way of DSG to limit framework variances and response time and integrate the nonbasic vertex and continuing nonbasic manner to increase energy productivity in order to fulfill short response time and greater energy proficiency. Assessing the short response time and high energy effectiveness objectives in film industry environments was not studied here. As they are simply based on point-by-point presumptions, concentrate on limiting energy utilization, or attempt to adjust energy and execution was also not analyzed in details. All the basic vertices just decide low response time on the basic way.

6. Conclusion and Future Direction

The work proposed a methodology for dealing with heterogeneous traffic-aware incoming data rate streams, Re-Storm at multiple traffic points, resulting in a short reaction time and great energy efficiency. It is divided into three parts, the first of which is a scientific model for fast response time and great energy efficiency. Then, distribution of resources considering DVFS methods presents effective optimal association methods and self-allocation of worker nodes. Furthermore, the results of the testing indicate that Re-Storm beats Storm by 20–30% for real-time streaming data utilized in Internet of things. It is not necessary to change any software approach or hardware device in this design; merely add an energy-efficient and traffic-aware algorithm. The design and development of this algorithm take into account all of the needs of the data generated by IoT applications. It has only fewer requirements to address a more complex big data problem, as well as an open-source platform.

The following research includes concerns for BDSC environment features, designs for huge real-time data streamed computing environments, influences on task topological graph with a cycle, and a dynamical extensibility of the various streaming data techniques may be studied further, developing BDSC platforms with high fault tolerance, scalability, throughput, and consistency for structuring such a system in a real-world BDSC context.

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.

Acknowledgments

References

[1] B. Albert, "Mining big data in real time," *Informatica*, vol. 37, no. 1, pp. 15–20, 2013.

[2] H. Mohanty, *Big data: an introduction*, p. 18, Springer, India, 2015.

[3] E. Benkhelifa, M. Abdel-Maguid, S. Ewenike, and D. Heatley, "The Internet of Things: the eco-system for sustainable growth," in *Proceedings of the IEEE/ACS 11th Int. Conf. on Computer Systems and Applications*, pp. 836–842, Doha, Qatar, November, 2014.

[4] J. Li, Z. Bao, and Z. Li, "Modeling demand response capability by internet data centers processing batch computing jobs," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 737–747, 2015.

[5] H. Shao, L. Rao, Z. Wang, X. Liu, Z. Wang, and K. Ren, "Optimal load balancing and energy cost management for internet data centers in deregulated electricity markets," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 10, pp. 2659–2669, 2014.

[6] P. Deepak, N. Surya, R. Rajiv, and C. Jinjun, "A dynamic prime number based efficient security mechanism for big sensing data streams," *Journal of Computer and System Sciences*, vol. 83, no. 1, pp. 22–42, 2016.

[7] H. Demirkan and D. Delen, "Leveraging the capabilities of service-oriented decision support systems: putting analytics and big data in cloud," *Decision Support Systems*, vol. 55, no. 1, pp. 412–421, 2013.

[8] L. Neumeyer and B. Robbins, "S4 : distributed stream computing platform," in *Proceedings of the IEEE Int. Conf. on*

Data Mining Workshops, pp. 170–177, Washington, DC, USA, October 2010.

[9] A. Ahmed and J. F. Naughton, "Static optimization of conjunctive queries with sliding windows over infinite streams," in *Proceedings of the ACM SIGMOD int. conf. on Management of data*, pp. 419–430, Paris, France, July, 2004.

[10] A. Amini, T. Y. Wah, and H. Saboohi, "On density-based data streams clustering algorithms: a survey," *Journal of Computer Science and Technology*, vol. 29, no. 1, pp. 116–141, 2014.

[11] K. Arun, K. Vamshikrishna, V. Kaladhar, and G. V. Prabhakara Rao, "CASH: context aware scheduler for Hadoop," in *Proceedings of the Int. Conf. on Advances in Computing, Communications and Informatics*, pp. 52–61, Chennai, India, September, 2012.

[12] C. Aggarwal, J. Han, J. Han, P. Yu, and P. S. Yu, "A framework for on demand classification of evolving data streams," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 5, pp. 577–589, 2006.

[13] J. Xu, Z. Chen, J. Tang, and S. Su, "T-storm: Traffic-aware Online Scheduling in Storm," in *Proceedings of the IEEE Int. Conf. on Distributed Computing Systems*, pp. 535–544, Madrid, Spain, June, 2014.

[14] S. Baskiyar and R. Abdel-Kader, "Energy aware DAG scheduling on heterogeneous systems," *Cluster Computing*, vol. 13, no. 4, pp. 373–383, 2010.

[15] C. H. Hsu, K. D. Slagter, S. C. Chen, and Y. C. Chung, "Optimizing energy consumption with task consolidation in clouds," *Information Sciences*, vol. 258, no. 8, pp. 452–462, 2014.

[16] D. Sun, G. Zhang, S. Yang, W. Zheng, S. U. Khan, and K. Li, *Re-stream: Realtime and Energy-efficient Resource Scheduling in Big Data Stream Computing EnvironmentsInformation Sciences*, vol. 319, pp. 92–112, 2015.

[17] S. Zhuravlev, J. C. Saez, S. Blagodurov, A. Fedorova, and M. Prieto, "Survey of energy-cognizant scheduling techniques," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1447–1464, 2013.

[18] Z. Zong, A. Manzanares, X. Ruan, and X. Qin, "EAD and PEBD: two energy- aware duplication scheduling algorithms for parallel tasks on homogeneous clusters," *IEEE Transactions on Computers*, vol. 60, no. 3, pp. 360–374, 2011.

[19] Z. Zhao, L. Zhuoyue, and L. Shurong, "Preemptive two-level priority real-time scheduling strategy for node system in wireless sensor network," in *Proceedings of the Int. Conf. on Advanced Infocomm Technology*, pp. 1–5, New York, USA, July, 2008.

[20] D. Sun, G. Fu, X. Liu, and H. Zhang, "Optimizing data stream graph for big data stream computing in cloud datacenter environments," *Int. J. of Advancements in Computing Technology*, vol. 6, no. 5, pp. 53–65, 2014.

[21] S. Gurmeet, K. Manku, and M. Rajeev, "Approximate Frequency Counts over Data Streams," in *Proceedings of the 28th ACM Int. Conf. On Very Large Data Bases*, pp. 346–357, Hong Kong, China, September, 2002.

[22] I. Keslassy, M. Kodialam, T. V. Lakshman, and D. Stiliadis, "On guaranteed smoothscheduling for input-queued switches," *IEEE/ACM Transactions on Networking*, vol. 13, no. 6, pp. 1364–1375, 2005.

[23] N. Kim, J. Cho, and E. Seo, "Energy-credit scheduler: an energy-aware virtual machine scheduler for cloud systems," *Future Generation Computer Systems*, vol. 32, no. 3, pp. 128–137, 2014.

[24] X. Liu, N. Iftikhar, and X. Xie, "Survey of real-time processing systems for big data," in *Proceedings of the 18th Int. Database Engineering and Applications Symposium*, pp. 356–361, New York, USA, July, 2014.

[25] M. Stonebraker, U. Çetintemel, and S. Zdonik, "The 8 requirements of realtime stream processing," *ACM SIGMOD Record*, vol. 34, no. 4, pp. 42–47, 2005.

[26] D. Pedro and G. Hulten, "Mining high-speed data streams," in *Proceedings of the ACM Int. Conf. On Knowledge Discovery and Data Mining*, pp. 71–80, Massachusetts, USA, August, 2000.

[27] K. Singh and R. Kaur, "Hadoop: addressing challenges of big data," in *Proceedings of the 2014 IEEE Int. Advance Computing Conf*, pp. 686–689, Navi Mumbai, India, February, 2014.

[28] G. Sudipto, "Tight results for clustering and summarizing data streams," in *Proceedings of the ACM Int. Conf. Proceeding Series*, pp. 268–275, St. Petersburg, Russia, March, 2009.

Retraction

Retracted: Smart Health Monitoring System with Wireless Networks to Detect Kidney Diseases

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. Dhanke, N. Rathee, M. S. Vinmathi, S. Janu Priya, S. Abidin, and M. Tesfamariam, "Smart Health Monitoring System with Wireless Networks to Detect Kidney Diseases," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3564482, 11 pages, 2022.

Research Article

Smart Health Monitoring System with Wireless Networks to Detect Kidney Diseases

Jyoti Dhanke,[1] Naveen Rathee,[2] M.S. Vinmathi,[3] S. Janu Priya,[4] Shafiqul Abidin,[5] and Mikiale Tesfamariam [6]

[1]*Department of Engineering Science (Mathematics), Bharati Vidyapeeth's College of Engineering Lavale, Pune 412115, Maharashtra, India*

[2]*Department of Electronics and Communication Engineering, IIMT College of Engineering, Greater Noida 201310, Uttar Pradesh, India*

[3]*Department of CSE, Panimalar Engineering College, Bangalore Trunk Road, Nazarethpet, Poonamallee, Chennai 600123, Tamilnadu, India*

[4]*Department of Electronics and Communication Engineering, K. Ramakrishnan College of Engineering, Samayapuram, Tiruchirappalli, Tamilnadu 621112, India*

[5]*Department of Computer Science, Aligarh Muslim University, Aligarh 202002, Uttar Pradesh, India*

[6]*Department of Software Engineering, College of Computing and Informatics, Haramaya University, Dire Dawa, Ethiopia*

Correspondence should be addressed to Mikiale Tesfamariam; mikiale.tesfamariam@haramaya.edu.et

It is essential to change health services from a hospital to a patient-centric platform since medical costs are steadily growing and new illnesses are emerging on a worldwide scale. This study provides an optimal decision support system based on the cloud and Internet of Things (IoT) for identifying Chronic Kidney Disease (CKD) to provide patients with efficient remote healthcare services. To identify the presence of medical data for CKD, the proposed technique uses an algorithm named Improved Simulated Annealing-Root Mean Square -Logistic Regression (ISA-RMS-LR). The four subprocesses that make up the proposed model are a collection of data, preprocessing, feature selection, and classification. The incorporation of Simulated Annealing (SA) during Feature Selection (FS) enhances the ISA-RMS-LR model's classifier outputs. Using the CKD benchmark dataset, the ISA-RMS-LR model's efficacy has been verified. According to the experimental findings, the proposed ISA-RMS-LR model effectively classifies patients with CKD, with high sensitivity at 99.46%, accuracy at 99.26%, Specificity at 98%, F-score at 99.63%, and kappa value at 98.29%. The proposed system has many benefits including the fast transmission of medical data to the medical personnel, real-time tracking, and registration condition of the patient through a medical record. Potential enhancement of the performance measures the provider system's hospital capacity and monitoring of a significant number of patients with a concentrated average delay.

1. Introduction

An IoT is a key concept that focuses on the modeling and connecting of Internet-linked physical objects using the power of computer systems. Instead of using a certain number of high-powered such as computers, tablet devices, mobile phones, and so on [1, 2], the IoT is largely used for a range of applications using a large number of low-power devices. Such as wrist bands, refrigerators, and umbrellas.

Combining IoT with Cloud Computing (CC) is beneficial when it comes to the creation of innovative approaches. An observation model has been built underneath the IoT integration and CC to monitor the state of patients and successfully gather the information, even in distant locations of greater assistance to medical professionals [3]. This model was created to watch the patient's condition. In several situations, the IoT technique is often kept with the assistance of the CC environment to improve efficacy in the

deployment of productive resources, data storage, and computational and processing capabilities [4].

The rate at which technological advancements occur is so rapid that almost anything may be connected over the World Wide Web [5]. The phrase IoT pertains to anything that is connected through the Internet. This term describes objects associated with the network and multiple objects connected to process information, and make decisions but also respond to the physical and virtual world [6]. IoT elements including cell phones, engines, sensors, and computing devices generate a significant quantity of data in real-time. The integration of cloud computing with the IoT is the greatest option. CC is indeed the fundamental virtual structure that can be utilized efficiently, and it encompasses many services like servers, networks, and storage [7]. Therefore, the combination of IoT and CC has substantial advantages and benefits in terms of algorithmic resource management, along with the existence and utilization of customizable resources and the usage of these resources as services [8]. There are a significant number of individuals throughout the globe who are afflicted with cardiovascular disease and diabetes may result in a stroke, heart, and kidney failure.

Blood pressure and heart rate sensors, together with an IoT-based cloud and Wi-Fi communications were used in the development of the system for monitoring the health of patients [9]. On the other hand, an actual electrocardiogram-based monitoring system is not included in the system. One such system has been proposed using an Arduino processor, the cloud, an ECG sensor, and a Raspberry [10]. Wi-Fi is used to make the connection to the Internet, and a cloud server receives the data, processes it, and then notifies the appropriate medical personnel in the event of an emergency [11]. It has been proposed that sensors equipment and a heartbeat Fitbit gadget are linked to the CPU through Bluetooth used to create a low-power health monitoring system. It is possible to use a Raspberry Pi as a server to store and process the vital signs of a patient in an emergency and send an alarm to the attending physician [12, 13]. CC is not supported on this system, and the user has a restricted range of motion to work within. Using temperature, pulse rate, ECG, and fall sensors and a Wi-Fi interface to transport the hospital's data to a cloud server, an IoT - based health system was built for real-time monitoring [14]. The system offers to monitor in real-time and makes a medical determination for the patient based on their current data and patient history. On the other hand, the system is not very precise when it comes to making judgments or, in the case of an emergency, alerting medical services [15]. The methods described in the aforementioned body of research provide several difficulties, including those about patient service area, the precision of decision-making, the database, warnings, real-time monitoring from any place, and the dimensions of the device [16].

2. Related Works

The purpose of developing and implementing smart watches wireless-based system for health monitoring that reads several vital parameters and then makes a selection based on a methodology that requested the server to contact the necessary person in the event of an emergency [17]. The system gathers information about the patient by employing several sensors and then transmits that information in real-time to a cloud server by way of an Arduino controller. The data is saved inside the cloud using an algorithm that was designed specifically to automatically monitor the health of the patients [18]. The system allows for a degree of adaptability for interaction between patients and the healthcare team using alerts, and it achieves a high level of precision in terms of decision-making [19]. When it comes to delivering medical treatment to a significant number of patients, healthcare institutions and caregivers face a variety of challenges that must be overcome. Because of the explosion in the use of contemporary biomedical sensors and IoT devices, an increasing number of sophisticated medical care and health monitoring systems have been created [20].

The majority of recent research has focused on the early identification of chronic disorders such as CKD, heart disease, and diabetes mellitus. In these studies, numerous factors that may influence chronic disease have been utilized. Despite this, the time complexity and performance of the illness prediction process continue to be significant issues [21]. This is even though all of the critical characteristics that are required for disease prediction are present. When attempting to diagnose CKD, the preliminary components often include a few influencing characteristics. These factors account for the long execution time and low precision of the prediction procedure. In addition to these qualities, however, additional symptoms such as insomnia, chest pain, nausea, and other indicators may also aid in the early diagnosis of probable CKD [22]. CKD stands for chronic kidney disease. Therefore, making use of the approach of feature selection may enhance the overall efficiency of CKD prediction made using a variety of classification strategies.

Then, we apply the important components derived from earlier research for CKD diagnosis in addition to the CKD severity level evaluation, which has not been explored before [23]. This was something that had not been done before. To evaluate the effectiveness of the categorization techniques that were put into practice, the samples needed for this study were taken from a selection of samples [24]. Also, in furthermore to the vital features that have been applied in recent research to predict CKD, several additional clinical symptoms are taken into consideration based on the proposed model [25]. These symptoms significantly contribute to the accuracy with which CKD and its severity degree can be predicted [26]. Therefore, to improve the implementation time and effectiveness of the testing and training process during the classification stage, we pertain to the influencing functionalities based on the clinical observations and experiences of physicians. The findings from earlier studies, provide three groups of crucial characteristics that have an impact on CKD [27]. In addition, if CKD is detected, this model can provide predictions on its subtype. The CKD level is the single most critical predictor in predicting CKD, although it has not been taken into account in any of the prior studies. The entire work done in the past was only devoted to making predictions about the presence of CKD, and the

samples were divided into two categories: healthy individuals and CKD patients.

The CC approach is more applicable for finding the basic linear function for NN to provide enough data training [28]. When a model is implemented, the classification procedure focuses on achieving the highest possible level of performance. To solve the problem that occurs in the Neural Network (NN)-CC technique's local optimum, modified techniques of NN-CC have indeed been developed and implemented. The main weights of the neuron connection are responsible for the operation of the NN, and the projected model makes use of the MCS scheme to bring down the overall Root Mean Square Error (RMSE) measurement which is calculated during the NN training process [29]. When contrasted with the NN-CS methodology, the obtained findings indicate that the NN-CC model has achieved an optimal level of function. The overall accuracy of CKD is brought down by the presence of missing values in a dataset. Because the conventional methods are executed during the data preparation stage, the data cleaning task is necessary to complete the work of filling in the missing data and getting rid of the scores that are not useful. It is anticipated that there will be a reevaluation technique used for all stages of CKD in which missing data have been calculated and inserted into the appropriate areas [30]. Even though conventional models are effective, a specialized illness detection system inside a healthcare dataset is still required to guarantee that CKD results are accurate. Within the context of this discussion, the FS task is seen as an essential component of the data categorization process.

This paper presents an ideal IoT and diagnosing chronic kidney disease (CKD) using a cloud-based decision support system, to deliver efficient virtual care to patients. The SA-oriented FS and Root Mean Square Propagation (RMSProp) driven Logistic Regression (LR) model termed ISA-RMS-LR is used by the proposed technique to identify the presence of CKD based on medical data. The proposed model has comprised a group of four different subprocesses, which include data collection, preprocessing, FS, and categorization. The incorporation of SA for FS contributes to the improvement of the classifier results generated by the ISA-RMS-LR model. Using a standard dataset for CKD, the validity of the ISA-RMS-LR model's predictive ability was investigated and confirmed. The results of the experiments suggest that the proposed ISA-RMS-LR model is superior to the methodologies that are being examined in terms of its ability to correctly diagnose and categorize CKD.

3. Proposed Model

Figure 1 shows the whole procedure for developing the proposed algorithm of the ISA-RMS-LR model. It depicts the data collecting process that occurs in a variety of ways. Following that, data preparation occurs, and this preprocessed data gets delivered into the ISA-FS model. The ISA-FS approach will choose the optimum subset of attributes, whereas the RMS-LR system will categories them.

In the first step, data is gathered via IoT devices that are connected to patients, a different benchmark healthcare dataset, individual health information, and a hospital administration server. The majority of the medical information provided by IoT devices linked to either a person is essential. A sensor attached to a person, for example, obtains accurate patient records at regular times [31]. CKD data were translated into an acceptable format in three processes. In the beginning, a format transforming operation is performed in which real data are transformed into a .csv file format. Several attribute values together in the dataset, like Absence, Present, and Yes or No, are simultaneously converted to numeric values such as 0 and 1. The median approach would then be used to complete a dataset's missing data.

Figure 2 depicts the general SA technique procedure. The physical aim is for molten material to cool at a minimum pace to assure the process of establishing thermal equilibrium at each temperature. When the temperature approaches absolute zero, the material achieves its ground state, which is composed of relatively low crystalline lattices. When a material solidifies at a temperature below its melting temperature, it forms an unsatisfactory framework that does not consume less energy [32]. In addition, Monte Carlo methods were used to predict the appearance of a material at a particular temperature. A little random disruption to the present state "l" of material with energy S_1 created the following state "n" with energy S_n. As long as S_n is below or equal to S_1, the condition may be considered current. Otherwise, the state has been spent with the following probability.

$$\mathrm{Exp}\left(\frac{(S_1 - S_n)}{(K_b T)}\right), \tag{1}$$

where T-current temperature; K_b-Boltzmann constant.

The Metropolis criteria are given the acceptance rule that was discussed, and the model was used in conjunction with the Metropolis criterion. A state of thermal equilibrium is reached by the material whenever it is heated to any degree because the temperature has indeed been progressively dropping.

$$P_T(X = 1) = \mathrm{Exp}\left(-\frac{S_1}{K_b T}\right)_n \sum_n \exp \frac{(-S_n)}{(K_b T)}, \tag{2}$$

where the sum of the energies of all possible states at temperature "T" serves as the denominator in equation (2).

3.1. Feature Selection Approach. To diagnose CKD, the proposed method uses the FS methodology. The success of the provided FS technique is contingent on the success of the ISA method which has seen widespread use for combinatorial optimizing tasks. It is presumed that the ISA model is one of the primary search approaches. This will allow the algorithm to find better ways to solve the problem. In this regard, the SA model has the quality of being able to achieve the best solution in contrast to the naive local search algorithm in every given set of conditions.
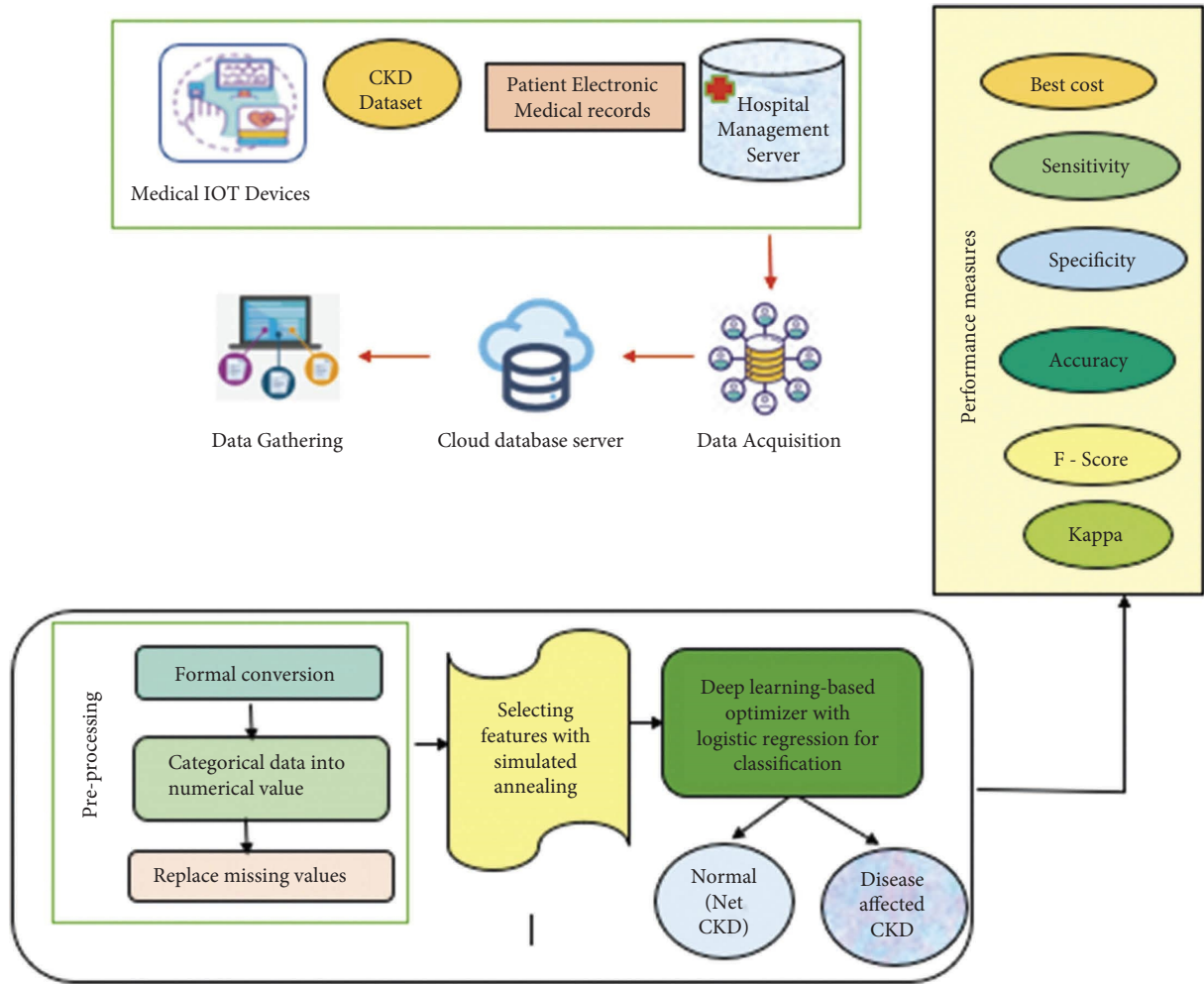
Figure 1: Proposed CKD system.

In most cases, it has been chosen in a completely arbitrary way because it has been a decision to be the best option. The use of the cost function allows for the estimation of the expense of an optimal state in a consecutive manner. When the temperature is unable to match the termination criterion, a solution that is adjacent to the most recent optimum solution is picked, and the cost is evaluated [33]. If the cost of a chosen adjacent solution is lower than or the same as the cost of the most recent optimum solution, then the most recent optimal solution may be swapped out for a new neighbor solution. An arbitrary function within the range has been chosen, even though the cost of nearby solutions is greater when contrasted with the cost of the present ideal solution (0, 1). When we attain this stage, the replacement of the optimum solution is initiated whenever the random value measured against is at its minimum $e = (\cos (Wn - \cos (e = \mathbf{W_a}))/T$. After the temperature has been lowered using equation (3), the process is repeated as many times as necessary until the temperature satisfies the termination conditions.

3.2. RL-Based Classification Model.
A classification task typically generates a presentation that is mapped with datasets of supplied classifications based on current data. Application of the deletion of required data content from a technique for determining the type of data. In the majority of instances, a variable has been required by the LR model for binary classification. This section defines LR approaches by focusing on two types of difficulties. This method's primary objective is to predict the existence of chronic kidney disease, which is typically confirmed using a binary classification methodology. In addition, LR approaches are often used to diagnose illnesses, identify multiple sclerosis, and classify health care data. The function of LR is to predict the presence or absence of CKD. Equation (3) demonstrates that the LR model is mostly reliant upon that linear regression technique:

$$\mathbf{R} = \boldsymbol{\alpha} + \boldsymbol{\beta}_1 \mathbf{i}_1 + \boldsymbol{\beta}_2 \mathbf{i}_2 + \cdots + \boldsymbol{\beta}_n \mathbf{i}_n. \tag{3}$$

Classification challenges are analogous to continuous-measurement-required linear regression problems. The
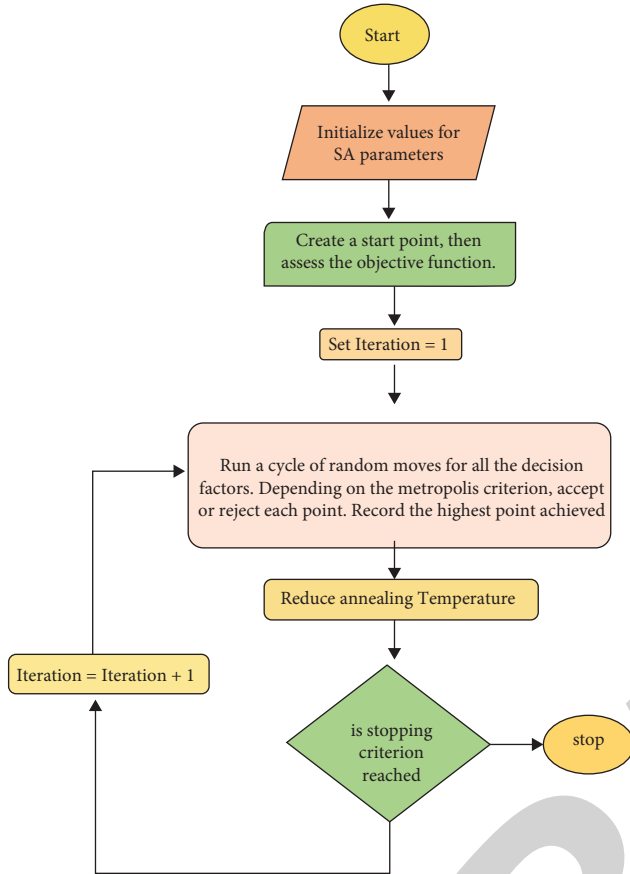
Figure 2: Simulated annealing model.

anticipated value of the categorization model falls between 0 and 1. The outcome is 1 if the numbers are above a specified threshold; otherwise, it is 0. Therefore, the LR parameter sequence falls inside the range [0, 1]. LR indicates that this level employs a sigmoid function. The entire linear model that detects the disease when such a sigmoid function has been applied is a feature.

$$Pr\,(A = +1|B) \sim \beta.B \text{ and}$$

$$Pr\,(A = -1|B0 = 1 - Pr\,(A = +1|Z),$$

$$\downarrow \sigma\,(x) = \frac{1}{1 + e^{-x}} \in [0, 1],$$

$$Pr\,(A = +1|B) \sim \sigma\,(\beta.B) \text{ and,}$$

$$Pr\,(A = -1|B) = -Pr\,(A = +1|B).$$

(4)

Classification is used at both the positive and the negative class levels here. Where "P" represents the existence of CKD, "Q" indicates the collection of eight elements' independent variables. Each regression model "Q" has a coefficient value represented by, which represents the weight. As determined by the LR model, these database values consist of weight values. It provides several linkages between Q and P for different weights. LR parameters may be altered to get an
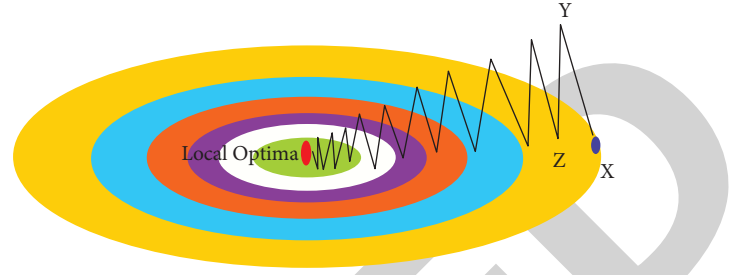
Figure 3: General RMSProp progress model.

optimal classification output. At this phase, the RMS model is utilized to choose the LR parameters.

The RMSProp model is based on the construction of such a weighted sum of grades, including such Gradient Descent (GD), using different progression parameters. Consider the scenario and evaluate a cost function made up of contours, with the red dot representing a local optimum. The very first iteration for GD begins at position 'A' and ends at point 'B;' Figure 3 shows the other end of an ellipse. The next phase of GD concludes at point 'C.' It reaches the optimum solution in ascending and descending dimensions faster than any prior iteration of GD. Although there is a greater learning value implementation, a lateral oscillation has a significant amplitude and decreases the GD while removing the higher learning value implementation. Vertical oscillations are caused by biases, but the horizontal motion is caused by weight. Vertical oscillation is reduced when the 'weights' are increased in value and the bias is adjusted. In backward propagation (BP), these parameters dW and db are utilized to modify W, whereas b is a fixed value.

$$W = W - \text{learningrate} * dW,$$
$$B = B - \text{learningrate} * db.$$

(5)

RMSprop uses progressively average values of either a squared of dW and db rather than the nonepoch-dependent distribution of dW and db.

$$SdW = \beta * SdW + (1-\beta) * dW^2,$$
$$Sdb = \beta * Sdb + (1-\beta) * db^2,$$

(6)

where β signifies an alternative hyperparameter that accepts the values 0 and 1, respectively. To compute the new weighted average, the weight from the average of previous values and the square of current values are assumed. In the factors updating process that follows the estimate of upper bound averages, SdW is quite small and Sdb is rather large.

$$W = W - \text{learning rate} * \frac{dW}{\sqrt{SdW}},$$
$$b = b - \text{learning rate} * \frac{dW}{\sqrt{Sdb}}.$$

(7)

TABLE 1: Dataset description.

Description	Value
Instances count	535
Features count	28
Class count	3
Percentage of positive samples	72.56%
Percentage of negative samples	37.52%
Data source	Own data from UCI data repository

TABLE 2: Analysis with CKD detection feature selection methodology results.

Methods	Best cost	Selected features
Proposed system	0.01054	7, 14,8,10,20,6,5,13,4,2,3
PSO-FS	0.03657	15,13,24,23,12,20,12,7,19,2,8,2,14,4, 1,5,18,17
GA-FS	0.03440	15,23,12,8,13,16,21,19,2,14,24,17,13, 5,3,9,3,20
PCA	0.04570	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19

3.3. Performance Validation. The effectiveness of the provided paradigm is validated by a comprehensive comparison with current practices. The proposed simulation was performed using MATLAB software. The number of Maximum iterations is set to 20, the Maximum Subiterations number is set to 5, and Initial Temp T0 is set to 10, with Temp. Reduction Speed alpha is set to 0.99. Table 1 depicts the dataset's overview and its accessible elements. The CKD dataset has a total of 400 occurrences with 24 characteristics. 250 occurrences are categorized as having chronic kidney disease (CKD), whereas the remaining 150 examples are designated as not having CKD.

4. Results and Discussion

The results of applying FS models to the CKD dataset are shown in Table 2, which can be found here. The optimum cost analysis of such SA-FS model that was supplied may be seen shown in Figure 4. With the best cost of 0.79, the table values proposed that perhaps the CFS model produced worse FS outcomes. In addition, it is shown that the best cost of the Principal Component Analysis technique is 0.04670, which is somewhat less than CFS but not significantly lower than competing models. The GA-FS and PSO-FS models have surpassed their predecessors and achieved almost equal best costs around 0.04640 and 0.04746, respectively. However, the proposed SA-FS model has selected just the thirteen qualities with the optimum cost, which is 0.01105. This lowest possible cost supplied by the SA-FS model ensured its greater performance in comparison to its predecessors.

Figure 5 depicts the various confusion matrices resulting from either the implementation of RMS-LR technology without FS at various epoch counts. The RMS-LR model has identified 182 cases as active and 150 instances as missing, as shown by the epoch count of 400. Likewise, for the epoch value of 800, the RMS-LR algorithm has categorized 117 cases as current and 150 occurrences as missing. Similarly, under the epoch value of 1200, the RMS-LR algorithm has identified 237 occurrences as present and 147 as missing. Concurrently, the RMS-LR algorithm has identified 242 cases as active and 146 occurrences as missing under the 2000 epoch count.

Figure 6 depicts several confusion matrices resulting from implementing ISA-RMS-LR using FS at varied epoch counts. Using such a different epoch count of 400 indicates that the ISA-RMS-LR method has identified 243 occurrences

FIGURE 4: Comprehensive cost estimate of the ISA-FS technique.

as present and 145 instances as missing in a collection. Similarly, a count of 800 epochs demonstrates that the ISA-RMS-LR system classifies 243 cases as present and 145 instances as missing. By this, it is seen that, under the epoch count of 1600, the ISA-RMS-LR method has categorized 247 occurrences that were present and 146 instances were absent. The ISA-RMS-LR algorithm has categorized 246 cases as present and 146 instances are absent under the 2000 epoch count.

After evaluating the classifier results provided by an ISA-RMS-LR system under various epoch counts, it was determined that perhaps the ISA-RMS-LR model may provide the most precise classification performance when the epoch count is between 1600 and 2000. Figure 7 demonstrates that in this specific round, the ISA-RMS-LR algorithm classifies 247 cases as present and 146 instances as absent.

Figure 8 illustrates the accuracy of the RMS-LR and ISA-RMS-LR models concerning the 2000 epoch data. The ISA-RMS-LR produces more accuracy than the RMS-LR method. The precision of both proposed models increases as the number of epochs increases, reaching a peak at 2000 epochs. Figure 9 illustrates the loss curves for the RMS-LR with ISA-RMS-LR algorithms for varied numbers of 2000 epochs. It has been demonstrated that the ISA-RMS-LR model produces a lower loss rate than the RMSPO-LR model. The real loss of the proposed models starts to decrease as the number of epochs increases and becomes constant around 2000 epochs. The addition of an SA-based FS procedure dramatically reduces the loss graph, as seen.
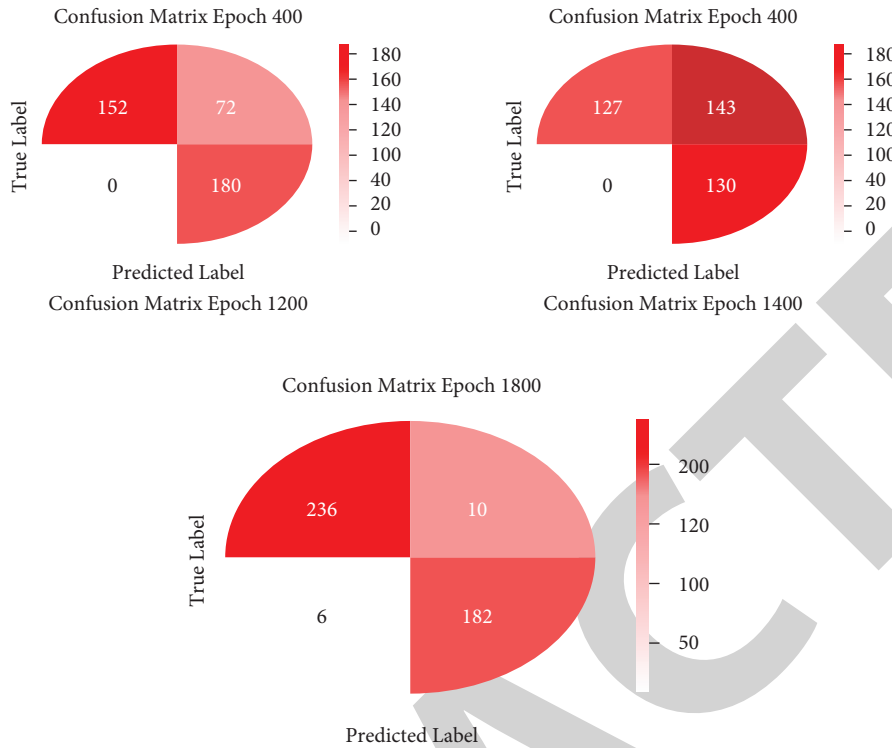
Figure 5: Confusion matrix for RMS-LR from 2000 Epoch FS.
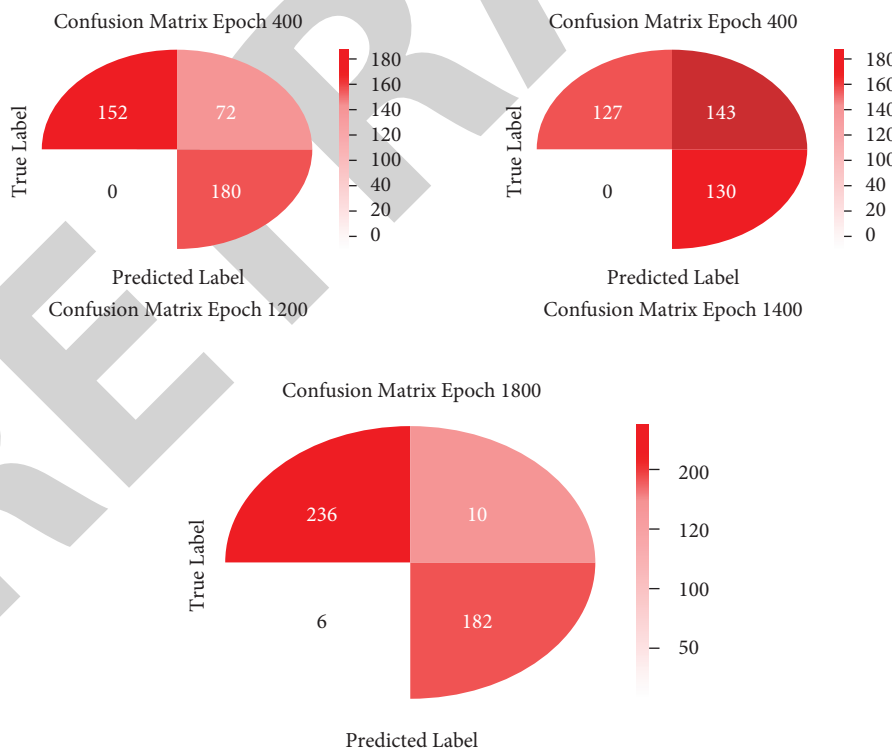
Figure 6: Confusion matrix of proposed system 2000 Epochs.

The outcomes of the ISA-RMS-LR model to those of current models for a variety of metrics compare in Table 3. Figure 10 presented a comparative examination of the sensitivity and specificity of the ISA-RSM-LR model's

findings. In addition, it is discovered that the D-ACO yields successful outcomes with sensitivity and specificity values around 86.00% and 83.35%, respectively. Upon continuation, the DT model yields a somewhat tolerable performance
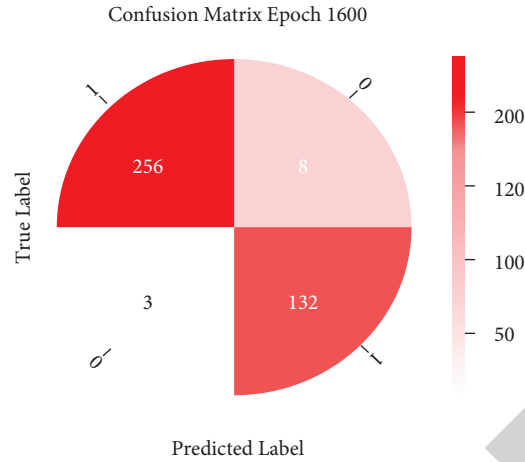
Figure 7: Confusion matrix for 1600th iteration.

(a)

(b)
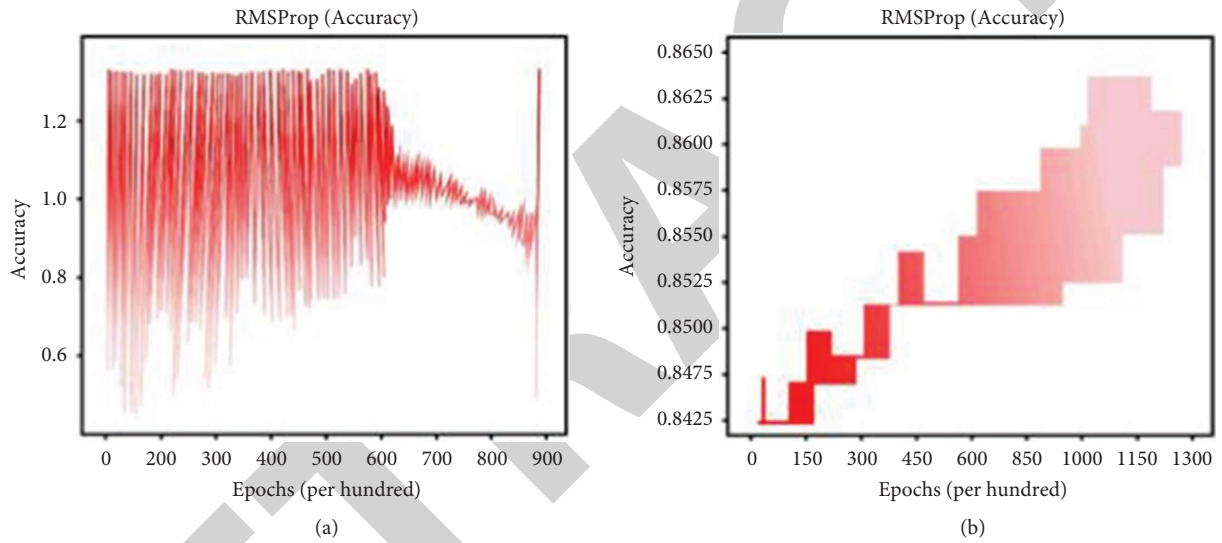
Figure 8: (a) RMSPO-LR accuracy graph with 2000 Epochs (b) ISA-RMS-LR accuracy graph with 2000 Epochs.

(a)

(b)
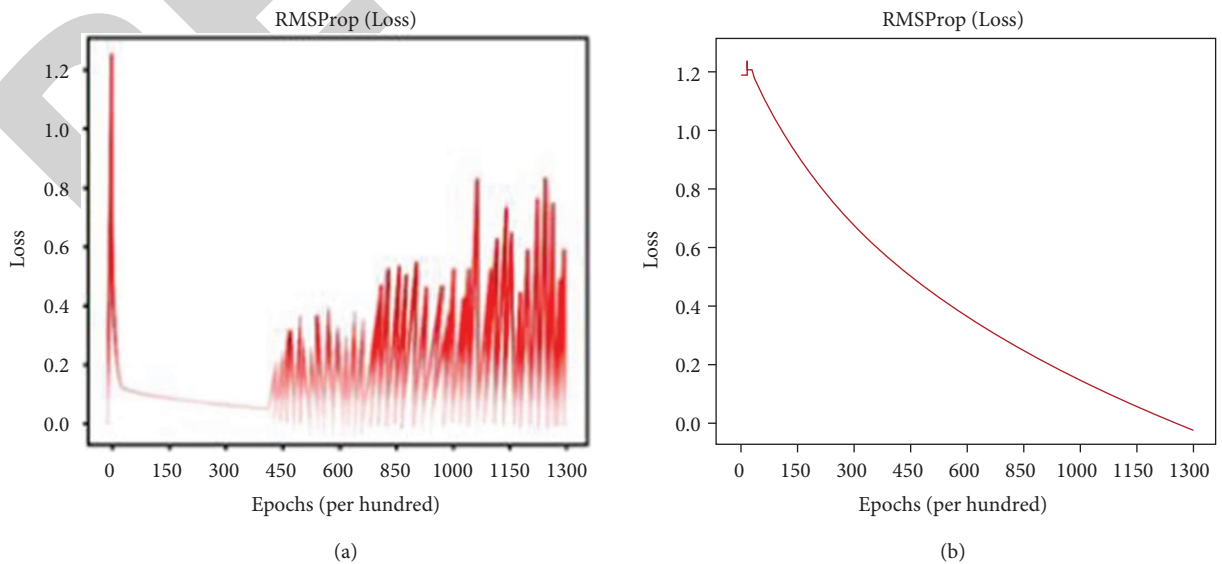
Figure 9: (a) RMS-LR Loss Graph with 2000 Epochs (b) ISA-RMS-LR Loss Graph with 2000 Epochs value.

Table 3: Proposed Method using the Performance Evaluation of CKD.

Classifiers	Performance measures				
	Sensitivity (%)	Specificity (%)	Accuracy (%)	F-score (%)	Kappa (%)
ISA-RMS-LR	99.46	98	99.26	99.63	98.29
RMSPO-LR	98.35	94.86	97.09	97.59	93.63
FNC	85.69	85.88	85.75	86.64	88.88
D-ACO	86.00	83.35	85.00	86.03	89.34

Input: CKD dataset with training
Output: FS with weight: W_a
  Step 1: Initialize $W_a \leftarrow$ NULL; $P \leftarrow 1000000$; $S \leftarrow 0.9$
  Step 2: It produces an initial solution, W_1;
  Step 3: Assign $W_a \leftarrow W_1$;
  Step 4: Calculate the initial solution cost as cost (W_a);
  Step 5: While $(T > 0.001)$ do
  {
  Step 6: Choose a nearby arbitrary solution, W_n from W_a that has one bit varied from W_a;
  Step 7: If $(\cost(W_a) = \cost(W_n))$;
  {
  Assign $W_a \leftarrow W_n$;
  }
  else
  {
  Develop an arbitrary number "a" range between (0.1);
  If $(a < e(\costt(w_n) - \cos t(W_a)/T))$.
  {
  Assign $W_a \leftarrow W_n$; $T \leftarrow S \times T$
  }
  }
  }

Algorithm 1: ISA-Based Feature Selection.
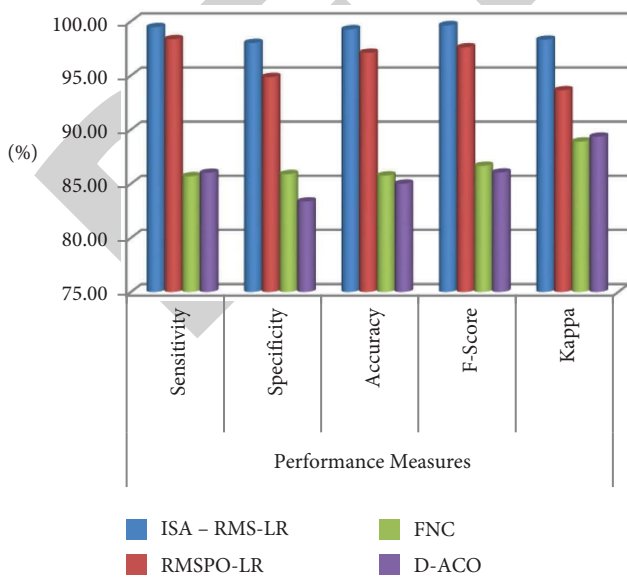
Figure 10: Performance measures.

using sensitivity and precision values of 91.38% and 88.28%, respectively. Similarly, the MLP model must achieve a classifier with sensitivity and precision values of around 93.30% and 93.86%, respectively. Simultaneously, the FNC model has shown consistent results, having sensitivity and accuracy values around 85.69% and 85.75%, respectively. The RMS-LR model has generated an effective classifier having sensitivity and accuracy values of around 98.35 percent and 97.09 percent, respectively. The maximal sensitivity and specificity values for the proposed ISA-RMS-LR models are 99.46% and 98%, respectively. In addition, it is shown that the LR has a significantly superior Kappa value of around 74.60%. In contrast, it is obvious that perhaps the XGBoost algorithm has achieved the same Kappa value at 75.42 percent. In addition, the PSO model beat current frameworks with a 68% higher Kappa value. In addition, ACO prefers to provide effective outcomes with a Kappa value of around 73.06%. However, with Kappa values of around 79.37%, the DT model produced slightly reasonable results. With Kappa values of 84.78 percent, the MLP model also achieves good classifier performance. Likewise, the FNC structure has shown consistency with a Kappa value of around 91.87 percent. Concurrently, it is obvious that the D-ACO approach produced a Kappa value that is even

higher, 88.33%. With both a Kappa value of 94.63 percent, the RMSPO-LR system is projected to provide competitive classification results. Therefore, the proposed ISA-RMS-LR method has produced results with a superior Kappa value of 93.26 percent. By evaluating the previous tables and graphs, it is determined that the RMSPO-LR model is a suitable medical diagnosis for CKD and may be utilized in real-time.

5. Conclusion

Using the ISA-RMS-LR model, this research assessed the optimal IoT and cloud-based decision support method for diagnosing CKD. Initially, the data collecting method captures the patient's information via medical devices. The gathered data are then preprocessed in preparation for future processing. The SA-FS method is then executed, and a subset of characteristics is provided to RMS-LR-based classifiers. The proposed classifier correctly identified the existence of CKD. The results of a detailed experimental investigation are verified using a benchmark dataset of CKD. The simulation results are reviewed for a range of epoch counts. According to the experimental findings, the proposed ISA-RMS-LR model effectively classifies patients with CKD, with high sensitivity at 99.46%, accuracy at 99.26%, Specificity at 98%, F-score at 99.63%, and kappa value at 98.29%. The obtained findings demonstrated the improved classification performance in comparison to the techniques evaluated. Clustering approaches may be used to enhance the performance of both the ISA-RMS-LR CKD diagnostic model as part of our future research. The framework intends to provide a complete patient history at home and around the world where the worldwide cloud server is frequently used, allowing the patient to easily access but also track health care providers from any location by logging through into platform, transferring past data, and accessing the web page data.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] B. Sumathy, S. Kavimullai, S. Shushmithaa, and S. S. Anusha, "Wearable non-invasive health monitoring device for elderly using IOT IOP conference series: materials science and engineering," *IOP Conference Series: Materials Science and Engineering*, vol. 1012, no. 1, Article ID 012011, 2021.

[2] Y. Xie, L. Lu, F. Gao et al., "Integration of artificial intelligence, blockchain, and wearable technology for chronic disease management: a new paradigm in smart healthcare," *Current Medical Science*, vol. 41, no. 6, pp. 1123–1133, 2021.

[3] G. Maniam, J. Sampe, R. Jaafar, and M. F. Ibrahim, "Smart monitoring system for chronic kidney disease patients based on fuzzy logic and IoT," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 2, 2022.

[4] H. Naz, R. Sharma, N. Sharma, and S. Ahuja, "IoT-inspired smart healthcare service for diagnosing remote patients with diabetes," in *Machine Learning for Edge Computing*, pp. 97–114, CRC Press, Boca Raton, FL, USA, 2022.

[5] S. Chitra and V. Jayalakshmi, "Analyze the medical threshold for chronic kidney diseases and cardio vascular diseases using internet of things," in *Proceedings of the 2021 4th International Conference on Computing And Communications Technologies (ICCCT)*, pp. 189–193, IEEE, Chennai, India, December 2021.

[6] S. Mishra, H. K. Thakkar, P. K. Mallick, P. Tiwari, and A. Alamri, "A sustainable IoHT based computationally intelligent healthcare monitoring system for lung cancer risk detection," *Sustainable Cities and Society*, vol. 72, Article ID 103079, 2021.

[7] L. T. Pugazhendhi, R. Kothandaraman, and B. Karnan, "Implementation of visual clustering strategy in self-organizing map for wear studies samples printed using FDM," *Traitement du Signal*, vol. 39, no. 2, pp. 531–539, 2022.

[8] P. Dileep and M. Thogaru, "Real time cloud computing based COVID-19 health monitoring system using IOT with integration of machine learning approach to create safety environment," *Annals of the Romanian Society for Cell Biology*, vol. 25, pp. 2076–2086, 2021.

[9] S. Maxwell and M. Grupac, "Virtual care technologies, wearable health monitoring sensors, and internet of medical things-based smart disease surveillance systems in the diagnosis and treatment of COVID-19 patients," *American Journal of Medical Research*, vol. 8, no. 2, pp. 118–131, 2021.

[10] J. Wang, Y. Wang, P. Pang, X. Jia, X. Yan, and Z. Lv, "Clinical analysis of the renal protective effect of GLP-1 on diabetic patients based on edge detection," *Journal of Healthcare Engineering*, vol. 2022, Article ID 6504006, 9 pages, 2022.

[11] R. Rajasekar and P. Sivakumar, "Swarm based intelligent transportation systems using internet of things in vehicular ad-HOC network," *Journal of Computational and Theoretical Nanoscience*, vol. 17, no. 12, pp. 5503–5508, 2020.

[12] R. P. Siguas, E. M. Solis, and H. M. Solis, "Design of a body temperature measurement system applied to health center personnel," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 2, pp. 20–28, Article ID 22502459, 2022.

[13] R. P. Siguas, E. M. Solis, H. M. Solis, and L. M. Zamudio, "Non-invasive glucose meter system for monitoring diabetic patients using the photoplethysmographic signal," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 6, pp. 14–19, Article ID 22502459, 2022.

[14] K. Sridharan and P. Sivakumar, "A systematic review on techniques of feature selection and classification for text mining," *International Journal of Business Information Systems*, vol. 28, no. 4, pp. 504–518, 2018.

[15] R. Perez-Siguas, E. Matta-Solis, H. Matta-Solis, and L. Matta-Zamudio, "Automated system for the classification of diabetic retinopathy," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 6, pp. 40–46, 2022.

[16] T. P. Latchoumi, R. Swathi, P. Vidyasri, and K. Balamurugan, "Develop new algorithm to improve safety on WMSN in health disease monitoring," in *Proceedings of the 2022 International Mobile and Embedded Technology Conference (MECON)*, pp. 357–362, Noida, India, March 2022.

[17] F. Shafqat, M. N. A. Khan, and S. Shafqat, "SmartHealth: IoT-enabled context-aware 5G ambient cloud platform," in *IoT in Healthcare and Ambient Assisted Living*, pp. 43–67, Springer, Singapore, 2021.

[18] A. S. Albahri, J. K. Alwan, Z. K. Taha et al., "IoT-based telemedicine for disease prevention and health promotion:

Hindawi

Research Article

Computed Tomography (Ct) Scan Assisted Machine Learning in the Management of Artifacts Related to Paranasal Sinuses and Anterior Cranial Fossa

Abdullah Musleh (iD)

King Khalid University, College of Medicine, Abha, Saudi Arabia

Correspondence should be addressed to Abdullah Musleh; amusleh@kku.edu.sa

Computed tomography (CT), through the use of ionizing radiation, allows us to assess the different parts of the body. It is made up of an X-ray tube that rotates rapidly around the patient generating the radiation beam. This is attenuated with the patient producing information, which is collected by the detectors that are opposite to the tube located in the gantry (part of the tomography equipment); finally, these collected data are sent to the computer that will reconstruct the information obtained and will represent it as an image on the monitor. In the practice of a study, artifices or artifacts may appear regardless of their origin, which limits the scan examination; this leads to stopping the examination and starting again, and added to this with the contrast media, they have to apply these drugs again. State-of-the-art scanners allow complete reconstructions to be performed with few projections, limiting radiation doses, by means of statistical algebraic reconstruction methods. The present work shows the simulation of artifacts in sinusitis diagnosis computed tomography images, the extraction of features from each image, and an automatic classification algorithm for the differentiation of artifacts. The results show that the algorithm is able to classify the simulated artifacts with a percentage of 90%.

1. Introduction

Computed tomography (CT), which employs ionizing radiation, enables us to evaluate various anatomical structures. The X-ray tube, which rapidly revolves around the patient and emits the radiation beam, is the main component [1]. This is diminished by the patient producing data, which is then gathered by detectors opposite the tube in the gantry (a component of the tomography equipment), and finally, these collected data are sent to the computer that will reconstruct the information obtained and will display it as an image on the monitor [2]. Sinusitis (Figure 1), a pathology marked by swelling and thickening of the mucosa lining the paranasal sinuses, may be brought on by obstructions, infections, or even anatomical variations [3]. Computed tomography is used frequently to diagnose sinusitis. It should be noted that sinusitis can affect both adults and children due to a viral or bacterial infection of the upper airways [4].

When we have an acute sinus infection, the tomographic study will reveal air-fluid levels because of the build-up of secretions that reduce the amount of air in the sinus and result in increased radiopacity and thickened mucosa. Chronic sinusitis is characterised by nonuniform, persistent radiopacity of the sinus, sclerosis, or thickening of the nearby bone, hypertrophy of the sinus mucosa, enlargement of the inferior turbinates, and nasal and sinus polyps [5]. The role of CT in this pathology makes it a tool that can be used directly in clinical practice because of the high sensitivity and specificity of contemporary technological equipment. Despite what has already been said, diagnosing the local sinusitis index is crucial. As a result, we are interested in figuring out how frequently patients have paranasal sinusitis (Figure 2), which can be diagnosed by CT scan [6].

The signs of acute sinusitis: Behind the skull, nose bridge and cheeks are air-filled cavities known as the sinuses. Ciliated cells that line the sinuses produce the mucus that
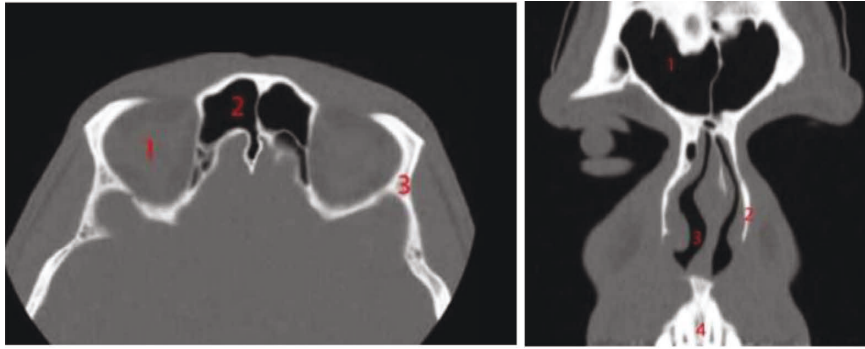
Frontal sinuses in axial and frontal section, computed tomography of paranasal sinuses.
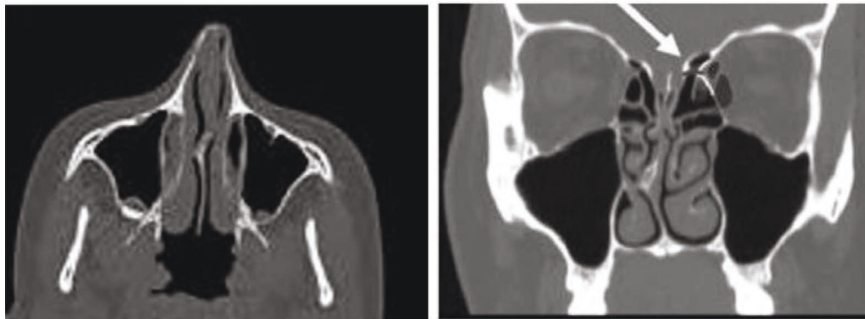
Maxillary sinuses in Axial and Coronal section.
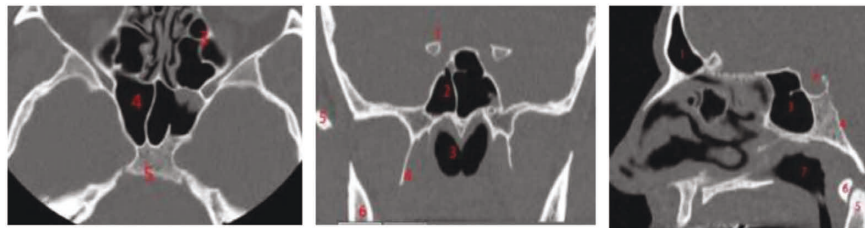
sphenoid sinuses in axial, frontal and sagittal slices computed tomography of paranasal sinuses
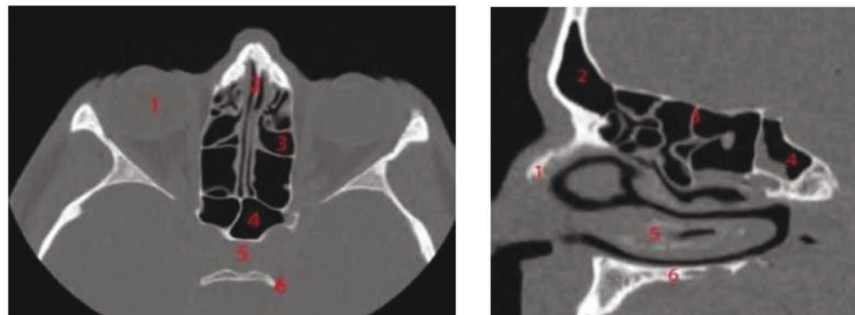
Ethmoidal cells in axial and sagittal section, computed tomography of paranasal sinuses

FIGURE 1: Various types of sinuses.

FIGURE 2: Tomography of the paranasal sinuses.

covers them. Bacteria and other contaminants are kept from spreading by mucus and nose hairs. A mucus that has thickened due to sinusitis can prevent the sinuses from receiving mucus. Sinus pressure rises as fluid accumulates. From bacterial growth, walls become infected. This affects 14% of Americans. In many cases, triggers include nasal or throat viruses or bacteria [7].

A sinus infection's early stages can cause symptoms mistaken for a cold, making diagnosis challenging. Congestion and fatigue can be brought on by either. A cold goes away in 5 to 7 days, but sinusitis can persist for weeks if left untreated. Flu-like symptoms, a fever, and facial pain are all symptoms of sinus infections. A quick office exam, patient history, and physical examination all help to identify acute sinusitis. In addition to inspecting your ears, nose, and throat, the doctor may also palpate your face to feel for sinus tenderness while asking you about the severity and duration of your symptoms. In order to make a diagnosis, your doctor might put a nasopharyngoscope through your nose. Doctors may order X-rays and CT scans to examine the paranasal sinuses and rule out any syndromes or other conditions. In most cases, sinus infections are treated and resolved in less than three weeks. Long-term sinus infections last longer than three weeks [8].

The mucosa lining the paranasal sinuses is closely related to the mucous membranes that line the nostrils and is completely pneumatized. The nasal cavity is encircled by them, and they drain into it. They are situated around the facial and cranial bones. It serves to warm, humidify, and protect the air that is inhaled as well as the nearby bones from which it gets its name. It also protects the lungs in times of trauma [9]. They consist of the sphenoid sinus, ethmoidal cells, and pairs of frontal and maxillary sinuses. They originate from the diverticula in the nostril wall. The superior, middle, and inferior turbinates are bony shell-shaped structures located inside the nostrils on the side walls. They have a mucosal lining that helps to maintain the quality of the air we breathe. The nasal septum, which is made of bone and cartilage on the backside and separates the right and left nostrils, is a bony structure [10]. The ostium connects them to the nostrils, and their pseudostratified cylindrical and ciliated respiratory epithelium lining ensures their physiology. However, if there is any obstruction, the partial pressures of oxygen and carbon dioxide inside the lungs become out of balance (14).

In computed tomography, the process of reconstructing images involves effectively utilising the interaction of x-rays with the body, taking into account the projections obtained in each beam path and accounting for the data recorded by each detector [11]. The same plane is subjected to numerous projections in order to reconstruct the cross-section of the body part being studied. In addition to the incidence and interaction of the X-rays, which are deterministic by the beam of the ray and their respective projections according to the angle of rotation and their arrival at the detectors, the human body is composed of structures of various densities that are very different from one another, allowing the creation of a recognition pattern that aids in the processing of images. As computed tomography (CT) technology has advanced, significant aspects of image acquisition should be considered. One of the most significant is the decrease in scan time, which is necessary to improve image quality and reduce exposure doses while also shortening reconstruction times. Although the number of projections can be determined in TAC, it is still not necessary to implement a mathematical formula for its processing; instead, effective algorithms are applied to evaluate it. Johann Radon developed one of the first image reconstruction methods in 1917 [12], who used a mathematical solution through a transform, which indicates that an image is determined by an infinite set of its projections [13].

In the mathematical processes of image reconstruction, the statistical methods fulfil a complementary function to the analytical methods but are differential at the same time, being the iteration methods our object of study in the present work. To ensure the best possible development of image processing, computed tomography relies on both mathematical and analytical data processing techniques. However, these techniques must take into account factors that make tomography a superior method of diagnosis in medicine, such as image quality, acquisition and processing times, minimization of radiation dose, reduction of noise and artifacts, and computational costs. Image quality is the result of the application of all the processes listed above. Still, in clinical practice, the patient's condition may act as a source of artifice generation and involuntary movements [14].

These aspects change the quality of the image and its diagnosis, necessitating a repeat exam and subjecting the patient to an excessive dose of radiation. In addition, when contrast media are provided, the travel times in the images are slowed down, which can also affect the diagnosis. Through an iterative reconstruction by approximation, which allows collecting and retrieving images to complete the study, it is possible to finish the study without needing to repeat the examination. Despite this, these methods require a complex algorithm and the appropriate implementation of the data for its processing. When an image appears with artifacts from a previous image, it is automatically recognised and moves on to carry out the correction process without the need to restart the study.

2. Theoretical Framework

2.1. Machine Learning.
Artificial intelligence is the branch of computing in which basic analog processes can be executed automatically in response to input with its respective output through programmed logic. Machine learning, as part of artificial intelligence, is responsible for generating algorithms that have the ability to learn, avoiding continuous programming; what is required is to feed the algorithm with data so that it learns and has autonomous decision alternatives. Machine learning is a scientific method that allows computers with the computational capacity to learn and extract patterns and correlate them by themselves. These patterns can then be used to predict behaviors that facilitate decision-making, and this is possible through information analysis called "training data" [15, 16].
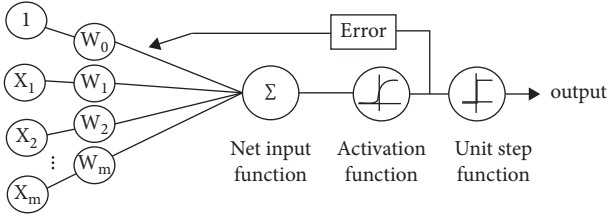
FIGURE 3: SVM architecture for regression taken from Acharya [20].

FIGURE 4: Linearly separable case [22].

2.2. Support Vector Machines (SVMs). The decision surface is learned by an SVM using two input classes. The data are mapped to a higher-dimensional feature space using a Gaussian kernel or another kind of kernel. The SVM theory is based on structural risk minimization, and neural networks have been used to highlight applications of the theory that have demonstrated outstanding performance [17]. The SVM resolves a quadratic problem where the number of inputs or training data equals the number of coefficients. Many applications exist for classification and regression, including image classification, character recognition, pattern classification, function identification. Another definition of the SVM is a system for efficiently training linear learning machines[18].

2.2.1. SVM for Regression. The goal of the SVM is to find a function $f(x)$ that has at most one deviation ϵ from the output y_i for all the training data, and at the same time, that is as small as possible. What has not been considered are minor errors ϵ, only those greater [19]. What is sought with the SVM for regression (Figure3) is to perform a mapping of the training data $x\epsilon X$, to a space of greater dimension F through a nonlinear mapping $\varphi: XF$, where the linear regression can be performed. The support vectors on which all the data that contain the most information possible depend [17].

2.2.2. SVM for Classification. To build a hyperplane that separates two classes, labelled $x, 1, +1$, so that the distance between the optimal hyperplane and the closest training pattern margin is the maximum [21]. Within the classification problems are the linearly separable and nonlinearly separable cases. Suppose we have S labeled points for training as shown in Figure 4.

2.3. Cross-Validation. Within the data validation techniques in machine training, it is obtained through theoretical physical models or with simulations through the hold-out validation methods and *k*-fold. When the amount of data for training and testing is limited, the hold-out method can be applied to estimate the error [23]; the objective is the adequate use of all the instances in D for the test training.

2.3.1. Leave-One-Out Cross-Validation. It is based on a confusion matrix where one record is used to test the validation classifier K times to run the training algorithm
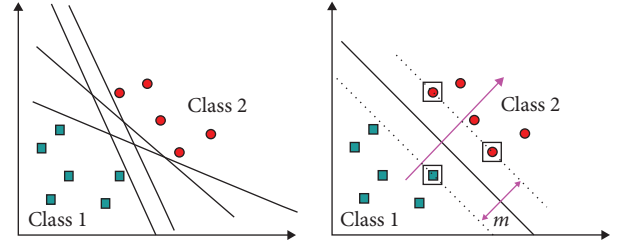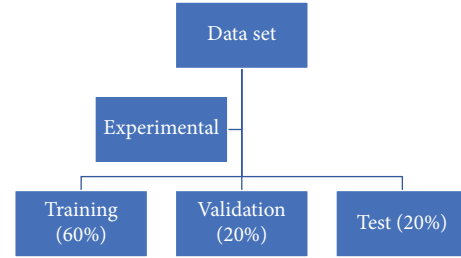
FIGURE 5: Hold-out separation.

through N iterations, excluding one from the sample process is repeated until leaving each of the samples outside the confusion matrix as shown in Figure 5.

2.3.2. Cross-Validation k Iterations. To perform *k*-folds validation, we split the original dataset into k equal parts; then, during model training, we use the first k of these parts as the test set and the remaining $(k − 1)$ parts as the training set. Each iteration of this procedure will collect data from a new test set [23].

2.4. ROC Curve. ROC curves are used to evaluate the performance of a classification method, which seeks to highlight event characteristics [24, 25]. The operating characteristics curve (receiver operating characteristics) represents the sensitivity and specificity for each threshold value and that allows comparing two or more classifiers based on their discriminant capacity, and we have the fraction of the true positive fraction (TPF) is plotted against the false positive fraction (FPF) which is given by the following function:

$$ROC_{(c)} = FTP_{(c)}, TPF_{(c)}; c \in -\infty, \infty. \tag{1}$$

One of the indices for estimating the efficacy of a system is the area under the ROC curve (AUC), whose value will always be between 0.5 and 1, which is why it is commonly used to compare the performance of the markers, the ROC provides a description of the separation between the distributions of positives and negatives.

3. Methodology

The methodology includes four steps in which a modification is made to a set of computed tomography images extracting features from said images using singular value decomposition, the training of a support vector machine

FIGURE 6: A proposed methodology for the simulation and classification of artifacts in CT images.

with different kernels, and validation to find the validity of the proposed algorithm as shown in Figure 6.

3.1. Image Set.
As input, a set of 43 paranasal sinus and skull base tomography images taken from different patients was used, where it was taken into account that the bone structures were differentiated. On the other hand, these images were found in the three body planes: sagittal, coronal, and axial.

3.2. Simulation of Artifacts.
Taking into account that there is no tomography equipment to carry out tests on images with noise or real artifacts, it was decided to apply different algorithms that allow simulating these effects on tomography images. In practice, there are two types of artifices that can be presented in computed axial tomography; depending on the origin, we can classify them as direct or indirect:

(i) Direct: they are generated by the patient: voluntary and involuntary movements

(ii) Metallic implants: osteosynthesis material, surgical clips

(iii) Indirect: detector misalignment; beam hardening; undersampling

3.2.1. Patient Movement.
In Euclidean space, a rotation is represented by a special kind of matrix called a rotation matrix. Rotation matrices in two dimensions look like the following equation (equation (2)):

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{2}$$

where θ represents the angle of rotation of the matrix.

3.2.2. Concentric Rings.
Two or more circles are said to be concentric if they have the same centre but different radii. Assuming a point (h, k) as the centre and a radius of r, the equation of the circle is

$$(x - h)^2 + (y - k)^2 = r^2. \tag{3}$$

Taking into account what is described in equation 31, rotation was used to generate different rings within the image with the same centre to simulate concentric rings' artifice.

3.2.3. Beam Hardening.
RGB values are converted to grayscale using the NTSC formula:

$$\text{GrayImage} = 0.299R + 0.587G + 0.114B. \tag{4}$$

The relative brightness of red, green, and blue light as perceived by the average person is well captured by this equation. The grayscale image was then converted to binary using Otsu's technique.

The threshold value is determined by Otsu's method in such a way that the spread within each segment is as small as possible while the spread between segments is as large as possible. To do this, the quotient between the two variances is calculated, and the maximum quotient is sought for a given threshold value. Let $p(g)$ be the probability of occurrence of the gray value $0 < g < G$ (G is the maximum gray value). Then, the probability of occurrence of the pixels in the two segments is

$$K_0: P_0(t) = \sum_{g=0}^{t} p(g), K_1: P_1(t) = \sum_{g=t+1}^{t} p(g) = 1 - P_0(t). \tag{5}$$

Once the image has been binarized, it is sought that the whites are attenuated to obscure the skull's internal structures. At this point, a threshold of 0.4 was used in order to detail such structures. Using the 43 original images, and taking into account that a transformation was applied to each of them in order to simulate the artifacts, a total of 172 images were obtained.

3.2.4. Feature Extraction.
Using the theory and considering the development, the null values of the matrix that are in the established range are discarded. When applying it to the 172 images, 172 matrices are obtained, and finally, the singular value of each element is obtained by passing each image through an algebraic method, which compresses the matrix that composes it. With the data obtained, a general matrix is fed that stores the result of the SVD of each image in order to train a machine learning algorithm in which interclass comparisons will be made with support vector machines and classification binary vectors.

3.3. Machine Learning.
The implementation of the algorithm to develop programs or solve specific problems can result in recognizable or repetitive patterns within the image, with which patterns or some kind of task can be predicted, improving its performance. For this case, a machine learning approach will be applied that allows to select and determine, and within a computerized axial tomography study bank, the images have some artifice and manage to classify them autonomously with respect to the original image. Finally, a cross-validation method of k-folds cross-validation is used in order to train the machine and find the real values of the simulation of the artifacts with respect to the reference image.

4. Results

Figure 7 shows the result of applying the rotation matrix to an image.
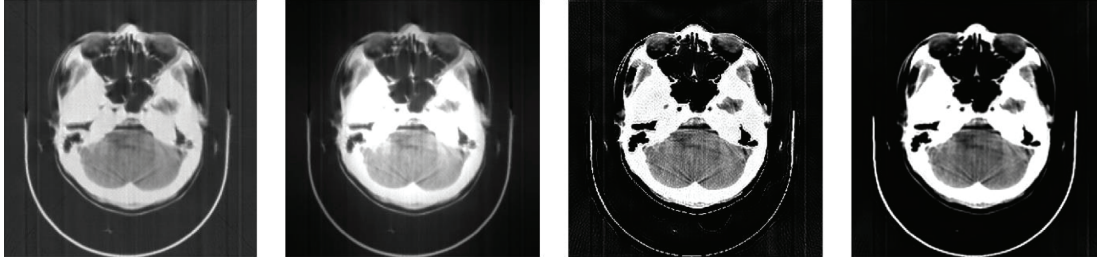
FIGURE 7: Orbit normal scan image, the same image with motion simulation.

On the other hand, from left to right, a normal brain scan image is followed by the simulation of the incident ray beam hardening device, that is, when there is a loss of contrast with respect to the original image due to the excessive absorption of the beam of the ray by the bony structures of the skull. Finally, you can recreate double artifices and concentric rings on the right, which considerably limits their evaluation.

In Figure 8(a), by means of the ROC curve, it can be seen that the AUC value (area under curve) is approximately 0.78, which indicates that this artifact has an error range of 12% with respect to the original images. This value was obtained through 30 iterations of the model, which can vary by 0.06 between iterations of the same cut.

It is important to mention that the samples or class must be of the same matrix size, which is ensured by the singular value decomposition (SVD); even so, the images correspond to different sections and brain structures, so the Kernel is of the Gaussian type.

In the case of cross-validation k-folds, between the original images and the images with beam hardening artifacts, the value is approximately 0.91 AUC, again demonstrating that it corresponds to a true positive; we can say that images are supported as a method of validation and classification. These were also performed in 30 iterations with a variation between each exercise and 0.09 between samples. The result can be seen in Figure 8(b). In the last case study, an AUC value of 0.63 (Figure 8(c)) was obtained, which is somewhat low compared to previous studies, but it is still considered a true positive, and this is due to although the rotated image maintains the same matrix size, the values of each composition pixel vary depending on the degree of rotation of the image. The result varies between 0.09 between the samples after 30 iterations.

With the values from the different samples, validations, and binary classifications between training samples and classes, it is possible to programme an algorithm that automatically identifies artifice. From this, machine learning's importance in image reconstruction as a method of learning and predicting them, optimising tomography imaging, can be seen (CT). As seen in Figures 8(a)–8(c), the trend of the tests is in the upper left corner, indicating that it has great prediction accuracy with each sample. We can also see that the model can distinguish between normal and fake images.

The device with the best test performance was the incident ray beam hardening with 83% efficiency, followed by the concentric ring device with 79% and the rotation device

with 73%. This is because the distortion of the original image varies depending on the artifice, with beam hardening causing the least distortion and rotation causing the most. Each artifice's simulation is based on a package of images. Due to the number of iterations and because they cannot be classified linearly, polynomial or Gaussian Kernel types are used when performing different tests. Even so, training the machine learning algorithm for image classification with CT artifacts is reliable. The case study will use a confusion matrix to redistribute the 172 images into 4 groups of 43 each. With these values, we can calculate the model's classification accuracy through the following formula:

$$\text{Precision} = \frac{TP}{(TP + FP)}, \tag{6}$$

where TP is True Positive, and FP is False Positive.

This means that 80% of the images will be recognised within or derived from the original image.

In the second validation measure, the Recall (Completeness) will be calculated, which refers to the prediction capacity of the machine learning model.

$$\text{Recall} = \frac{FN}{(TP + FN)}, \tag{7}$$

where TP is True Positive, and FN is False Negative.

This means that the model is able to recognise 90% of the images with artifice and the normal images.

Finally, the F-score or $F1$ will be calculated to determine the positive predictive value, and the performance of the model is the following equation:

$$F - \text{Score} = \frac{\text{Precision} \times \text{Recall}}{(Pres + Recall) \times 2}. \tag{8}$$

This indicates that the quality of the model is approximately 84% reliable, taking into account the entire sample, that is, the 172 images. This can be seen in the graphs of each ROC curve of the artifacts versus the original images, where each score is close to 1 as a true positive. What proceeds next is to carry out the same exercise, but with each of the proposed tricks against the set of original images to obtain the following Figure 9.

The data are taken in each of the cases, and it is evident that the percentages are relatively consistent in comparison with those obtained through the classification model of prediction proposed in MATLAB; on the other hand, for each artifact, it can be concluded that the artifact that most

(a)

(b)

(c)
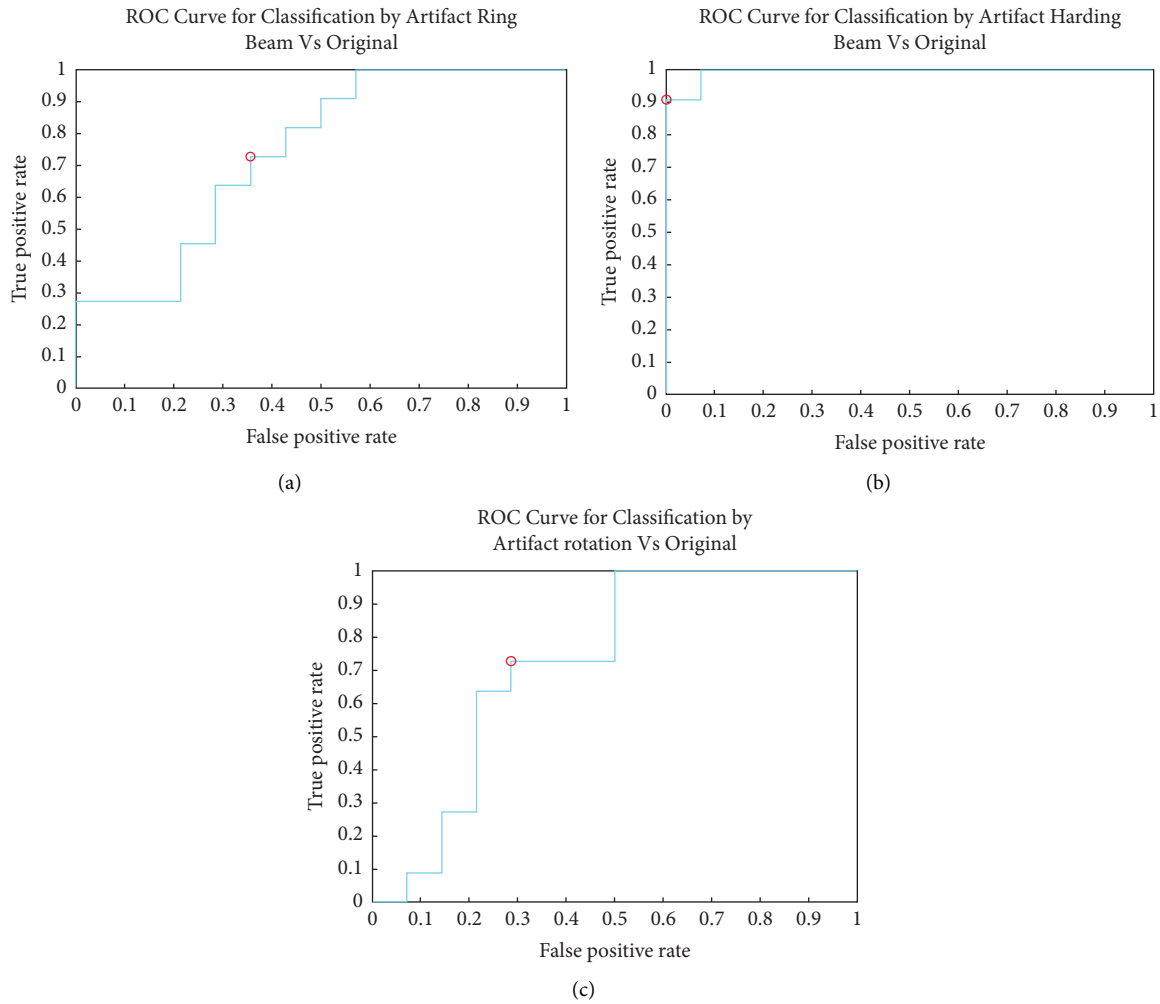
FIGURE 8: (a) ROC curve original image vs. concentric rings; (b) ROC curve original image vs. artifact hardening beam; (c) ROC curve original image vs. artifact rotation.

closely resembles the original image is hardening of the incident beam since the original image is maintained, but the intensity of the grayscale is affected, showing overexposure in the brain structures is consequently poorly defined. The second artifice with greater adherence to the original image is that of concentric rings, which preserves the original format but the distortion due to overlapping is evident in the brain structures. Lastly, the images with a rotation device are the ones that lose the most similarity with the original images, where the evaluation of any brain structure is totally limited.

5. Discussion of Results

The image quality is improved by the algebraic methods' significant reduction of all types of noise and artifacts. Another benefit of this method is that reconstruction can be completed without requiring all of the conventional projections of a typical study, which also results in a significant decrease in exposure doses per study. For example, when using iterative algorithms, matrices are recalculated in all possible directions to obtain new sinograms. When using
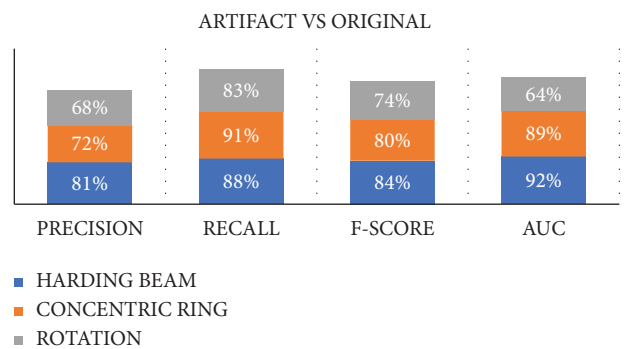
FIGURE 9: Final results of the classification algorithm with an interclass comparison between the simulated artifacts and the original image.

other statistical methods, corrections are made in the geometry using the incident ray beam approach, which also requires more computational work and, as a result, has a higher GPU cost [20–22]. From the foregoing, we can infer the followings:

(i) High contrast and superior quality images

(ii) Requires fewer projection data for image reconstruction

(iii) Considerable reduction in exposure dose

(iv) Longer reconstruction time

(v) High computational costs

Hybrid architectures that use both reconstruction methods and parallel algorithms are being developed to create more efficient reconstruction mechanisms. Higher image resolution means a larger initial matrix and more equations per pixel, which equals more tomograph detectors. Parallel algorithms and filtering techniques such as Gaussian, median, Wiener, and bilateral filters are being implemented to improve data acquisition efficiency [21].

As we have discussed, several ways to "clean" an image efficiently with statistical methods require complex adaptation to current hardware systems. It works with GPU systems, a graphic processing unit, to speed up iterative data processing, implement new algorithms, and achieve maximum reduction [17, 20, 21]. Statistical methods have the advantage of not requiring a large number of projections and data to reconstruct later. Noise and artifacts in images are not only caused by external factors, such as voltage arcs, improper instrument calibration, and misaligned detectors. Image quality improves with analytical and mathematical algorithms.

6. Conclusions

Any artifice chosen for the study can be duplicated from any scanned image, replicating a real medical image while testing the visualisation of sinusitis's important aspects. MATLAB could mimic the most common CT artifacts without changing the original matrix's dimensions. Based on the algebraic properties of each matrix, which can be transformed and used in other matrix operations, it was possible to duplicate tomography artifacts. Medical image classification algorithms are not linear due to matrix features when using normal and verification photos as training references. Artificially produced images require a Gaussian or polynomial Kernel to solve and represent. These photos can't be arranged on a normal plane for classification; hence, a linear solution isn't possible. The Gaussian Kernel allows features derived from artifact-filled simulated images to be consistent with the binary classification space defined; i.e., inside the classification group, we have two alternative labels between 0 and 1 to define positive or negative.

Using this binary classification, we can show how true positives and false positives behave in the sample. Machine learning helps computed axial tomography equipment's image reconstruction processes automatically recognise artifices and perform the appropriate modifications to lessen or eliminate image distortion. K-folds fit SVM classification data better, allowing model validation with an AUC (ROC). Cross-validation (K-folds) assesses the estimator's performance to determine the algorithm's prediction effectiveness. Iterations are used to retrain the model. Beam hardening

artifact drastically altered the image, allowing for sensitive classification. This shows that the binarization algorithm's threshold affects image disturbance.

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] O. Duliu, "Computer axial tomography in geosciences: an overview," *Earth-Science Reviews*, vol. 48, no. 4, pp. 265–281, 1999.

[2] A. Faisal, S. Parveen, S. Badsha, H. Sarwar, A. W. Reza, and W. Ahmed, "Computer assisted diagnostic system in tumor radiography," *Journal of Medical Systems*, vol. 37, no. 3, 9938 pages, 2013.

[3] M. Gumusok, M. Kaya, F. Cankal, N. Apaydin, and I. Tekdemir, "Role of anatomic variations of paranasal sinuses on the prevalence of sinusitis: computed tomography findings of 350 patients," *Nigerian Journal of Clinical Practice*, vol. 20, no. 11, 1481 pages, 2017.

[4] G. K. Scadding, "Rhinitis and sinusitis," *Clinical Respiratory Medicine*, pp. 409–423, 2008.

[5] I. Sumaily, I. Alarifi, A. Alahmari, M. Aloulah, and S. Alsaleh, "Sphenoid sinus involvement in chronic rhinosinusitis without polyps," *Allergy & Rhinology*, vol. 11, Article ID 215265672093447, 2020.

[6] M. Vestin Fredriksson, Kuoljok, Jenny, L. Flygare, D. Berggren, and K. Tano, "Clinical manifestations and symptoms of maxillary sinusitis of odontogenic origin demonstrated by cone beam computed tomography," *Journal of General Practice*, vol. 14, 2019.

[7] M. Badr, S. Al-Otaibi, N. Alturki, and T. Abir, "Deep learning-based networks for detecting anomalies in chest X-rays," *BioMed Research International*, vol. 2022, Article ID 7833516, 10 pages, 2022.

[8] N. Shaikh, A. Hoberman, D. H. Kearney et al., "Signs and symptoms that differentiate acute sinusitis from viral upper respiratory tract infection," *The Pediatric Infectious Disease Journal*, vol. 32, no. 10, pp. 1061–1065, 2013.

[9] R. Pourmousa, R. Dadashzadeh, F. Ahangarkani, and M. S. Rezai, "Frequency of bacterial agents isolated from patients with chronic sinusitis in northern Iran," *Global Journal of Health Science*, vol. 8, no. 5, p. 239, 2015.

[10] S. Nair, "Correlation between symptoms and radiological findings in patients of chronic rhinosinusitis: a modified radiological typing system," *Rhinology*, vol. 47, no. 2, pp. 181–186, 2009.

[11] T. K. Kim and J. Jeong, "Deviated nose: physiological and pathological changes of the nasal cavity," *Archives of Plastic Surgery*, vol. 47, no. 06, pp. 505–515, 2020.

[12] F. Casali, "Chapter 2 X-ray and neutron digital radiography and computed tomography for cultural heritage," *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage*, vol. 1, pp. 41–123, 2006.

[13] L. Yu, X. Liu, S. Leng et al., "Radiation dose reduction in computed tomography: techniques and future perspective," *Imaging in Medicine*, vol. 1, no. 1, pp. 65–84, 2009.

[14] M. Maray, M. Alghamdi, and M. B. Alazzam, "Diagnosing cancer using IOT and machine learning methods," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9896490, 2022.

[15] R. Booij, R. P. Budde, M. L. Dijkshoorn, M. van Straten, and Marcel van Straten, "Technological developments of X-ray computed tomography over half a century: user's influence on protocol optimization," *European Journal of Radiology*, vol. 131, Article ID 109261, 2020.

[16] J. Iwanaga, T. Tanaka, S. Ibaragi et al., "Revisiting major anatomical risk factors of maxillary sinus lift and soft tissue graft harvesting for dental implant surgeons," *Surgical and Radiologic Anatomy*, vol. 42, no. 9, pp. 1025–1031, 2020.

[17] A. Nuseir, A. Nuseir, M. Alsmirat, M. Al-Ayyoub, M. Mahdi, and H. Al-Balas, "Performance of different machine learning methods for sinus diseases classification," in *Proceedings of the 2021 12th International Conference on Information and Communication Systems (ICICS)*, Valencia, Spain, May 2021.

[18] Z. Rustam, N. Angie, J. Pandelaki, and R. E. Yunus, "Acute sinusitis classification using support and fuzzy support vector machines," *Journal of Physics: Conference Series*, vol. 1490, no. 1, Article ID 012029, 2020.

[19] Z. Li, X. Feng, Z. Wu, C. Yang, B. Bai, and Q. Yang, "Classification of atrial fibrillation recurrence based on a convolution neural network with SVM architecture," *IEEE Access*, vol. 18, Article ID 2920900, 1 pages, 2019.

[20] G. A. Betancourt, "Support vector machines (svms)," *Scientia et Technica*, vol. 1, no. 27, 2005.

[21] J. Aurelia, Z. Rustam, A. Laeli, and F. Maulidina, *Neural Network-Support Vector Machine for Sinusitis Classification*, pp. 1185–1189, 2020.

[22] T. Acharya, "Regional scale landslide hazard assessment using machine learning methods in Nepal," Edited by T. Yang, Ed., Thesis for: Doctor in EngineeringAdvisor, , Thesis for: Doctor in EngineeringAdvisor, 2018.

[23] P. A. Aguirre, *Study of machine learning techniques applied to crop classification in satellite images*, PhD thesis, National University of Cuyo, Mendoza, Argentina, 2019.

[24] O. S. Ahmed, E. E. Omer, S. Z. Alshawwa, M. B. Alazzam, and R. A. Khan, "Approaches to federated computing for the protection of patient privacy and security using medical applications," *Applied Bionics and Biomechanics*, vol. 2022, Article ID 1201339, 2022.

[25] S. Habib and Mona, *Improving Scalability of Support Vector Machines for Biomedical Named Entity Recognition*, A dissertation submitted to the Graduate Faculty of the University of Colorado at Colorado Springs, Colorado Springs, 2022.

Research Article

Vehicle Driving Risk Prediction Model by Reverse Artificial Intelligence Neural Network

Huizhe Ding [ID],[1,2] Raja Ariffin Raja Ghazilla [ID],[1] Ramesh Singh Kuldip Singh [ID],[1] and Lina Wei [ID][2]

[1]*Centre of Product Design and Manufacturing, Department of Mechanical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia*
[2]*School of Mining and Geomatics, Hebei University of Engineering (Hebei), No. 19 Taiji Road, Handan Economic and Technological Development District, Handan 056038, China*

Correspondence should be addressed to Raja Ariffin Raja Ghazilla; r_ariffin@um.edu.my

The popularity of private cars has brought great convenience to citizens' travel. However, the number of private cars in society is increasing yearly, and the traffic pressure on the road is also increasing. The number of traffic accidents is increasing yearly, and the vast majority are caused by small private cars. Therefore, it is necessary to improve the traffic safety awareness of drivers and help car manufacturers to design traffic risk prediction systems. The Backpropagation neural network (BPNN) algorithm is used as the technical basis, combined with the MATLAB operation program, to simulate the driving process of the car. Dynamic predictive models are built to predict and analyze vehicle safety risks. Multiple experiments found that: (1) in various simulations, the simulation driving process of MATLAB is more in line with the actual car driving process; (2) the error between BPNN and the actual driving prediction is within 0.4, which can meet the actual needs. Predictive models are optimized to deploy and predict in various traffic situations. The model can effectively prompt risk accidents, reduce the probability of traffic accidents, provide a certain degree of protection for the lives of drivers and passengers, and significantly improve the safety of traffic roads.

1. Introduction

With the continuous development of social science and technology, automobiles play a major role in traffic travel. Traffic accidents have caused widespread concern, seriously threatening the safety of human life and property [1]. Statistics show that about 1.2 million people are killed in traffic accidents every year around the world. Driver factors cause about 80% of tragedies. Bad driving behavior will cause great discomfort for passengers and even cause psychological pressure on drivers [2]. The driver needs to process the information in time while receiving the traffic information while driving, and the driving risk is everywhere. Safe driving requires drivers to maintain a vigilant attitude in driving [3].

There are numerous studies on the relationship between drivers and traffic accidents, but few studies on the risk warning of the vehicle itself. According to the basic physiological and psychological characteristics of drivers, combined with the specific functions of the car, the prediction of traffic risk is an emerging topic [4]. An artificial neural network (ANN) is a multilevel early-warning model, which uses data analysis technology to predict driver behavior and system, showing strong flexibility and adaptability [5]. The BPNN can perform correlation analysis on multiple complex relationships, make macro predictions of accidents, reduce the risk of accidents, and provide security [6].

First, a simulation driver system based on the MATLAB computer program is proposed. Dynamic simulation models are built to perform dynamic analysis of vehicle driving. Second, the BPNN is used to build different structures and analyze driving safety risk and accident prediction. The driver's association with accident occurrence is explored. Finally, combined with the error backpropagation

algorithm, different experimental data are detected, recorded, and compared. Driving safety risks and effective countermeasures are analyzed. This work intends to study the relevant factors of car accidents during driving, reduce the risk index of driving safety in people's travel and traffic, give early warning of possible accident risks, and conduct research. It is committed to reducing and preventing driving errors from the root cause and effectively improving the status quo of road traffic safety.

2. Materials and Methods

2.1. Realization and Simulation of Driving Simulation in MATLAB. Today, in the automotive industry, it is planned to use computer simulation to integrate more real system data, various machine plate components, and electronic control units into the vehicle model to develop and operate various types of automotive products. Research in this area has become one of the hot spots [7]. The operation of the internal electronic control system and mechanical dynamics of the automobile requires the use of related technologies to simulate the transient motion process of the automobile. The actual motion state process of the car is shown in Figure 1.

2.1.1. Simulation Environment. MATLAB software has been widely used in the design and simulation of control systems. At this stage, the field of simulation technology is a relatively mature scientific computing language, which is good at efficiently processing large amounts of data and visualizing data [8]. MATLAB is used to simulate the driving process of the car and analyze how the car engine and clutch, and other components break and operate in an emergency. The control system is simulated, and the risk factors affecting the driving control system of the vehicle are obtained more intuitively, which provides a strong scientific basis for the risk estimation and performance improvement of the vehicle in the future. MATLAB software is used to simulate the car driving system. The frame line of the car simulation driving system is shown in Figure 2.

Figure 2 is the frame line of the car simulation driving system. First, we open the window of the car simulation driving system and input various parameter values required for simulated driving, including the number of gears of the transmission, the number of front-wheel rotation angles, and the degree of throttle opening. [9]. The system will match the corresponding actual road conditions on the ground according to the different scenarios selected and import it into the dynamic simulation system of the car as the parameters to be matched later. After matching, the system performs dynamic simulation and presents the results with virtual technology. When different scenarios are selected to be simulated again, only the parameters of the car driving system need to be changed [10].

2.1.2. Mathematical Model of the Algorithm. In the vehicle driving speed control system, the position transmission belongs to the speed manipulation tool and is the main input part of the system. Its functions mainly include reasonable
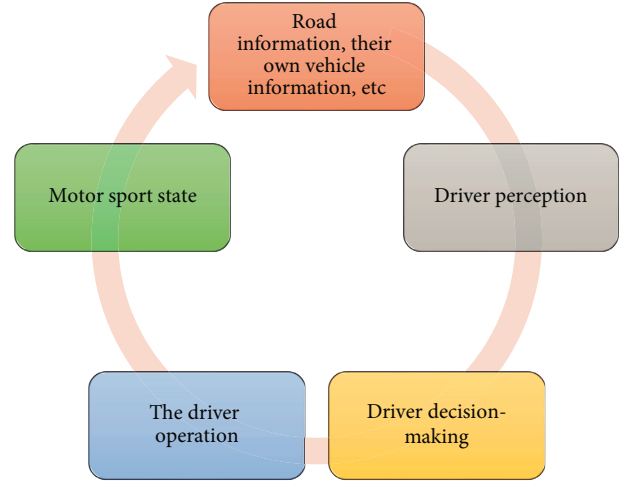
FIGURE 1: The flow of the motion state of the car.
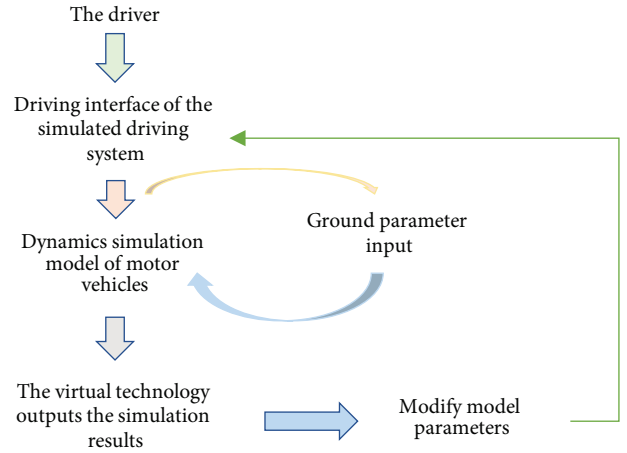
FIGURE 2: The simulated driving frame of the car.

control of the speed and transition of car driving, as shown in the following:

$$v = ax + b, x \in [0, 1], \tag{1}$$

x represents the position of the speed control mechanism; v represents the driving speed corresponding to it; and a and b represent constants.

In the vehicle control system, proportion-integral-derivative (PID) belongs to the core part, which mainly compares the current speed of the car with the speed at which it is specified. The difference between the two speeds is the current vehicle traction [11]. The calculation of the driving controller is shown in equations (2)-(4)as follows:

$$\text{Calculus equation part: } x(n) = x(n-1) + u(n), \tag{2}$$

$$\text{Differential equation part: } d(n) = u(n) - u(n-1). \tag{3}$$

TABLE 1: Basic parameter values.

Road slope	Adhesion coefficient	Initial speed n	Throttle opening	Steering wheel angle	Brake percentage	Gear
0.05	0.8	66.7	0.7	0.0	0.0	2
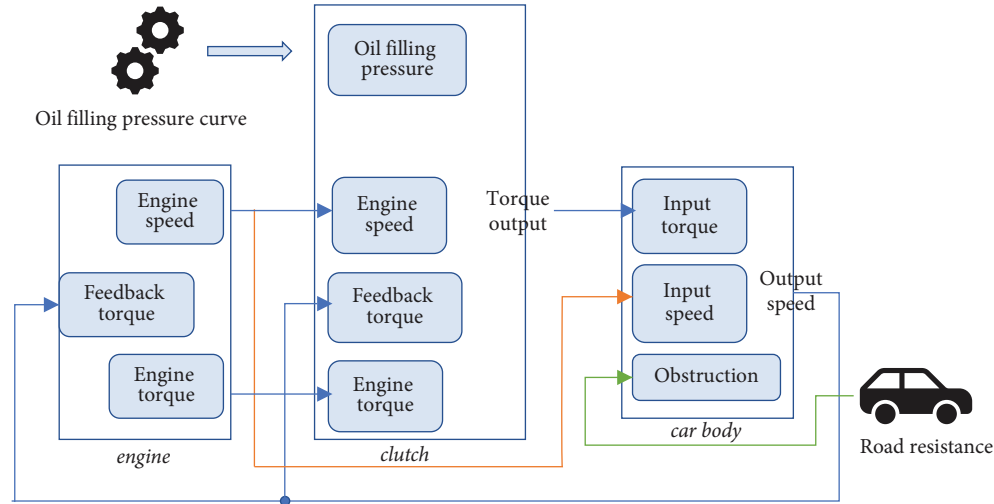
FIGURE 3: Simulation model of car driving.

The output section of the system:

$$y(n)Pu(n) + Ix(n) + Dd(n), \qquad (4)$$

$u(n)$ represents the system input value, that is, the difference between the current speed of the car and the specified speed; $y(n)$ represents the system output value, representing the traction force of the car; and $x(n)$ represents the state of the system.

In the vehicle driving control system, the vehicle power mechanism belongs to the actuator. Its function mainly uses the traction force to change the vehicle's speed so that the speed is the same as the specified value. The mathematical expression between traction and speed is $F = mv + bv$. F represents the traction force generated by the car; m represents the mass of the car; v represents the speed of the car; and b represents the drag factor of the car [12].

2.1.3. Parameter Setting. After opening the MATLAB interface, we enter the relevant parameter values in the GUI interface. Simulink's default simulation time is 10 seconds. In the initial stage of the engine, there is a certain initial speed, which is set to 2000 n/min [13]. The car is running at a constant speed at this stage. We accelerate and apply the brake while the car is driving at a constant speed. The specific parameter value settings are shown in Table 1.

According to the operation of the actual vehicle starting process, MATLAB simulation software is used to establish the simulation model of vehicle starting. The model includes four plates: engine, transmission, body, and road resistance plates [14]. According to the modeling idea of the simulation model, the complex molecules in the system are subdivided into smaller plates. The vehicle driving simulation model is shown in Figure 3.

In Figure 3, after the simulation model is built, the simulation parameters are set according to the system requirements. The position transmission part of the speed control mechanism adopts the Slider Gain module. It can limit the range of the input signal x of the position transmission. Among them, the parameter minimum value is set to 0, the maximum value is set to 1, and the initial value is set to 0.55. In the Gain module, the gain value is set to 50. In the Constant 1 module, the constant value is set to 45. In the vehicle powertrain subsystem model, the Gain module is adopted, its value is 1/m, and the specific value is 1/1000. The value of the Gain 1 module is b/m, and the specific value is 20/1000. The automotive powertrain adopts the Integrator integral module. The initial state is set to 0, which is the initial speed value of the car. In the driving controller subsystem model, the Delay block is used to implement the PID controller, the initial state is set to 0, and the sampling time is set to 0.02 s. The simulation time range of the system is set from 0 to 500 s. The variable-step continuous solver is selected. The rest of the other modules and simulation parameters use MATLAB default values [15].

3. ANN

Based on the basic characteristics of artificial intelligence to simulate the brain of living things, a new network model ANN is proposed. It can simply simulate and simplify different biological networks and provide help for the research of various functions of the real biological brain nervous system. The basic components of ANN include

neurons and synapses. Among them, the function of a neuron is called a processing unit, which can use digital language to express the information processing method of the biological neuron, formally describe the biological neuron, and simulate the function and function of the expression structure [16].

3.1. Structure of an Artificial Neuron.

An artificial neuron is an abstract concept of a biological neuron. The transmission of signals between neurons is also a transmission of values. Multiple signals are received and then converted into a new round of signals to be transmitted, which belongs to a nonlinear structure. The specific structure is shown in Figure 4.

Figure 4 shows the structural model of an artificial neuron, which is seen as a simple processor that performs a weighted summation of incoming signals. The specific description is shown in the following:

$$Y = W_0 + W_1 X_1 + W_2 X_2 + \ldots + W_m X_m. \tag{5}$$

In equation (5), $(X1)$, $(X2)$, ..., X_m represent the input value; W_1, W_2, ..., W_m represent the weights corresponding to the input value, that is, the synapse of the simulated neuron; Y represents the output of the neuron; and W_0 represents the system bias.

The network composed of artificial neurons is ANN. Since the emergence of this theory, it has been continuously improved, and now a relatively complete theory has been formed, which is widely used in different fields [17]. The application research of neural networks in traffic safety risk warning is based on the practical application background of the deep learning theory, combined with the actual driving situation, analyzes the possibility of vehicle risk in some cases, and carries out early warning according to the operation mechanism of the neural network, so as to reduce the possibility and harm of traffic accidents [18].

3.2. Establishment of the Risk Assessment Model.

There have been many attempts at applying ANN in risk assessment. Many practices have confirmed that it is a better risk assessment method and can be used as a supplement to traditional risk assessment. The system can analyze the risk assessment with the law of actual risk occurrence and provide a basis for the early-warning system. An ANN risk assessment model is proposed, as shown in Figure 5.

Based on ANN, a risk model for driving safety is established. The practical application is divided into six steps. (1) The number of layers in the middle hidden layer of the neural network is determined, including the number of input, output, and hidden layer nodes. (2) Characteristic parameters and state parameters are determined. In the process of risk analysis, the internal structure and external parameters in the risk assessment system are determined, and it is ensured that the results and state characteristics of the input parameters can be correctly reflected. (3) The neural network system needs to provide a variety of learning samples for feedback learning during execution and analyze the network system for parameter values in different states.
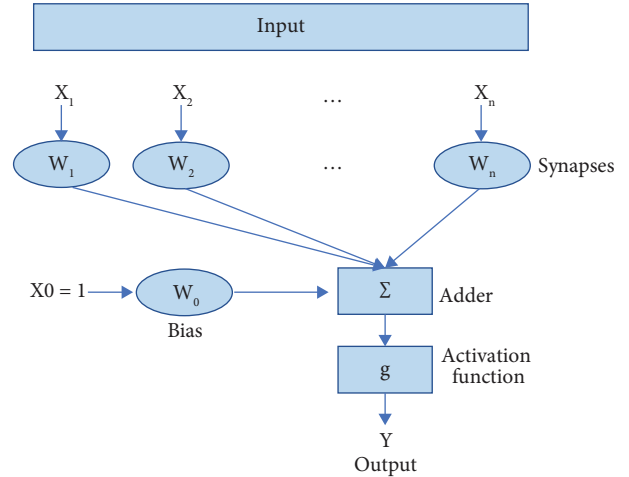
FIGURE 4: The nonlinear structure of an artificial neuron.

This process is that the neural network determines the connection weights and errors of the network system according to the selected samples. (4) The nonlinear sigmoid function is chosen as the action function. (5) Network learning is used to confirm the structure of the neural network to establish a knowledge base for risk analysis of the nervous system so that the evaluation system has certain inference and prediction functions. (6) The actual driving situation is analyzed. The actual eigenvalues of the system that have been calculated are input into the neural network with a prediction function. After the data are processed by the internal risk assessment system, the evaluation results are obtained. Additionally, this result is fed into the neural network as a new computational sample, which enriches the risk assessment system database [19].

4. BPNN Model

4.1. The Principle of BPNN.

The BPNN is a multilayer feedforward network trained according to the error backpropagation algorithm, and it is one of the most widely used neural network models. It can learn and store many input-output relative relationships without revealing the mathematical equations describing this mapping in advance. The learning method of BPNN uses the steepest descent method and backpropagation to repeatedly modify the weights and thresholds of the system to minimize the sum of squared errors [20]. The BPNN structure includes three parts: input, hidden, and output layers. The specific structure is shown in Figure 6.

In Figure 6, the operation process of BPNN consists of two parts: forward propagation of data and backpropagation of error. During forward propagation, data are passed in from the input layer. After passing through the hidden layer, it is passed out from the output layer. If the actual output situation of the output layer is different from the ideal output, the result turns to the backpropagation stage of the error [21]. The backpropagation of the error means that the output error enters the hidden layer in a certain way and is transmitted back to the input layer, and the error data are apportioned to all units of each layer to obtain the error data of each layer unit. These error data are used as the basis for
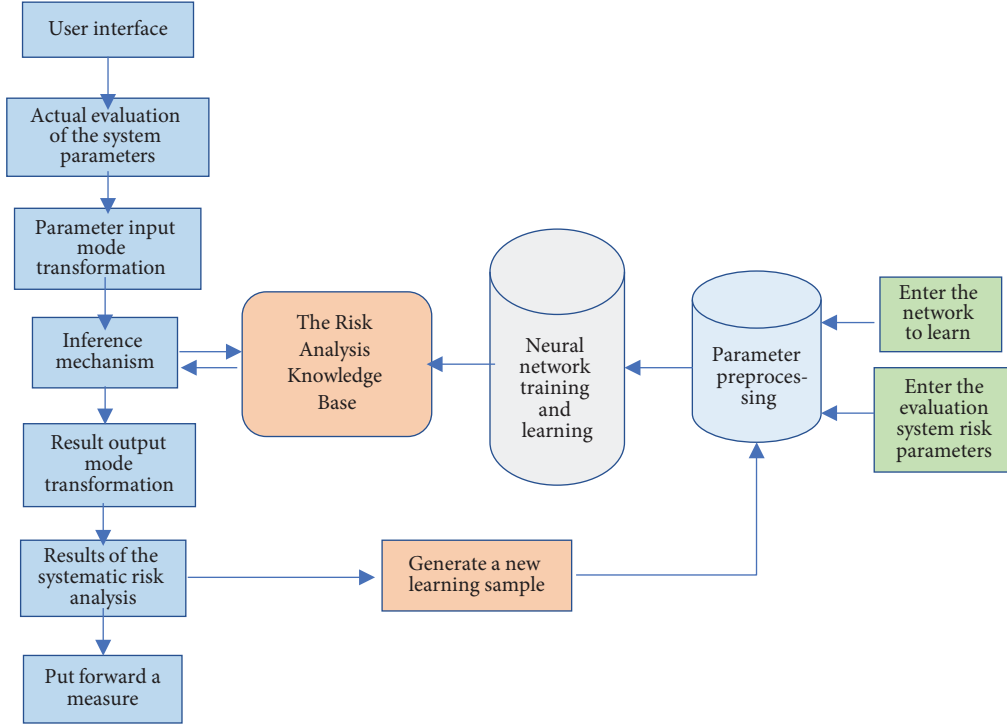
Figure 5: ANN's risk assessment model.
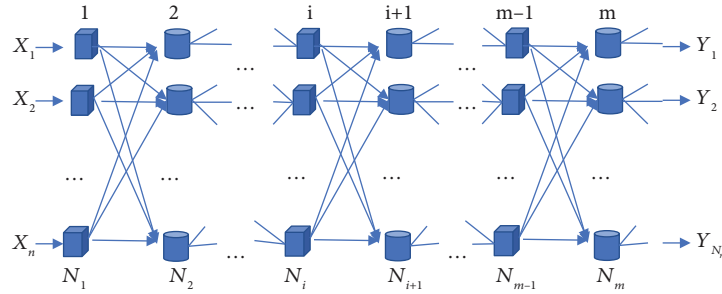
Figure 6: The structure of BPNN.

correcting the weights of each unit [22]. The BPNN consists of input, output, and hidden layers. N_1 is the input layer, N_m is the output layer, and the rest are hidden layers.

4.2. Algorithm Derivation of Backpropagation Three-Layer Neural Network. During the forward propagation, the activation function of neuron j produces the induced local domain $V_j(n)$ at the input, as shown in the following:

$$v_j(n) = \sum_{i=0}^{m} w_{ji}(n) y_i(n). \qquad (6)$$

$\varnothing_j$ is the activation function. The representation of the function signal $y_j(n)$ at the output of neuron j is shown in the following:

$$y_j(n) = \varphi_j\big(v_j(n)\big). \qquad (7)$$

In the process of error backpropagation, $y_j(n)$ represents the actual output of neuron j; $d_j(n)$ represents the expected output of neuron j, which is the j-th element of the expected response vector $d(n)$. The error signal $e_j(n)$ is shown in the following:

$$e_j(n) = d_j(n) - y_j(n). \qquad (8)$$

Minimizing the root mean square error makes the function continuously differentiable. Here, the instantaneous error energy of neuron j is given, as shown in the following:

$$E_j(n) = \frac{1}{2} e_j^2(n). \qquad (9)$$

The error energies of all output layer neurons are summed up, i.e., the sum of the instantaneous error energies of the entire network is obtained, as shown in the following:

$$E_n = \sum_{j \in C} E_j(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n). \tag{10}$$

In equation (10), set C refers to all neurons in the output layer.

The backpropagation algorithm minimizes E_n by modifying the weights multiple times. The iterative gradient descent method is used to apply a correction value $\Delta w_{ji}(n)$ to the synaptic weights $w_{ij}(n)$. It is proportional to the partial derivative $\delta E(n)/\delta w_{ji}(n)$. According to the differential chain method, this gradient is given by the following:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)}. \tag{11}$$

The partial derivative $\delta E(n)/\delta w_{ji}(n)$ represents a sensitive factor, which determines the search direction of the synaptic weight $w_{ji}(n)$ in the weight space.

$$\frac{\partial E(n)}{\partial e_j(n)} = e_j(n), \tag{12}$$

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1, \tag{13}$$

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi_j'(v_j(n)), \tag{14}$$

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n). \tag{15}$$

Equations (7)-(10) are substituted into equation (6), as shown in the following equations (16) and (17):

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n)\varphi_j'(v_j(n))y_i(n), \tag{16}$$

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)}. \tag{17}$$

$\Delta w_{ji}(n)$ is the correction function of $w_{ji}(n)$; η represents the learning rate of error backpropagation; the negative sign represents gradient descent. Equation (11) is substituted into equation (12), as shown in the following:

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n). \tag{18}$$

In equation (18), $\delta_j(n)$ is the local gradient defined according to the delta law. It indicates the required changes in synaptic weights, as shown in the following:

$$\delta_j(n) = -\frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} = e_j(n)\varphi_j'(v_j(n)). \tag{19}$$

4.3. Algorithm Flow of BPNN. The backpropagation three-layer neural network has been recognized by the public as

the most suitable model for simulating input, output, and early warning [23]. The specific operation process is shown in Figure 7.

In Figure 7, the algorithm operation process includes forward and backward propagations. During forward propagation, information data enter from the input layer and are processed by the hidden unit to the output layer. The state of a neuron in each layer only affects the state of neurons in the layer below it [24]. If the output layer does not get the desired result after a series of operations, the data are transmitted back to the backpropagation, and the error signal is returned along the original neuron connection path. When returning, the weights of the neuron connections will be modified one by one [25]. This process is iteratively processed. Finally, the signal error value is within the allowable range. Eventually, the output value is close to the desired output.

5. Tips for Traffic Risk Information

There may be latent risk information in the static and dynamic information of traffic. Information that may lead to risk events is risk information. In China's urban road environment, with mixed and nonmotor vehicles, the risk information is mainly motor vehicles, nonmotor vehicles, and pedestrians. There may be latent risk information in the static and dynamic information of traffic. If the response to avoid the risk information is not made in time, it will lead to risk events. Risk events can lead to traffic conflicts, and traffic conflicts can lead to traffic accidents, as shown in Figure 8.

In Figure 8, when the risk information appears, the driver's physiological and psychological factors significantly impact the reaction speed, including the influence of factors such as age, gender, personality, and education level. In addition, the driver's sensitivity to risk information, judgment ability, and feedback speed during driving are also related to the type of risk information. During the driving process, people or things that affect the normal driving trajectory and driving speed of the vehicle can become risk information. In addition to common people and vehicles, there may also be sediment, falling rocks, animals, and poultry on various mountain roads or country trails. The risk information mainly includes pedestrians, nonmotor vehicles, and motor vehicles.

The risk warning model must be consistent with the level of human security risk perception before it can be accepted in practical applications. During the driving process, the driver subjectively feels the changes in the information of the road segment and has a stress response to the information of hidden risks, which is a comprehensive performance. For example, when a driver suddenly encounters a pedestrian crossing the road in front of him, illegally climbing over the guardrail, or the car in front of him suddenly braking, suddenly changing lanes, and so on in normal driving, the driver needs to pay attention and discover the risk information in time in advance. The driver's skills can respond to the predicted risk information and take actions such as braking or steering in advance to avoid dangerous situations.
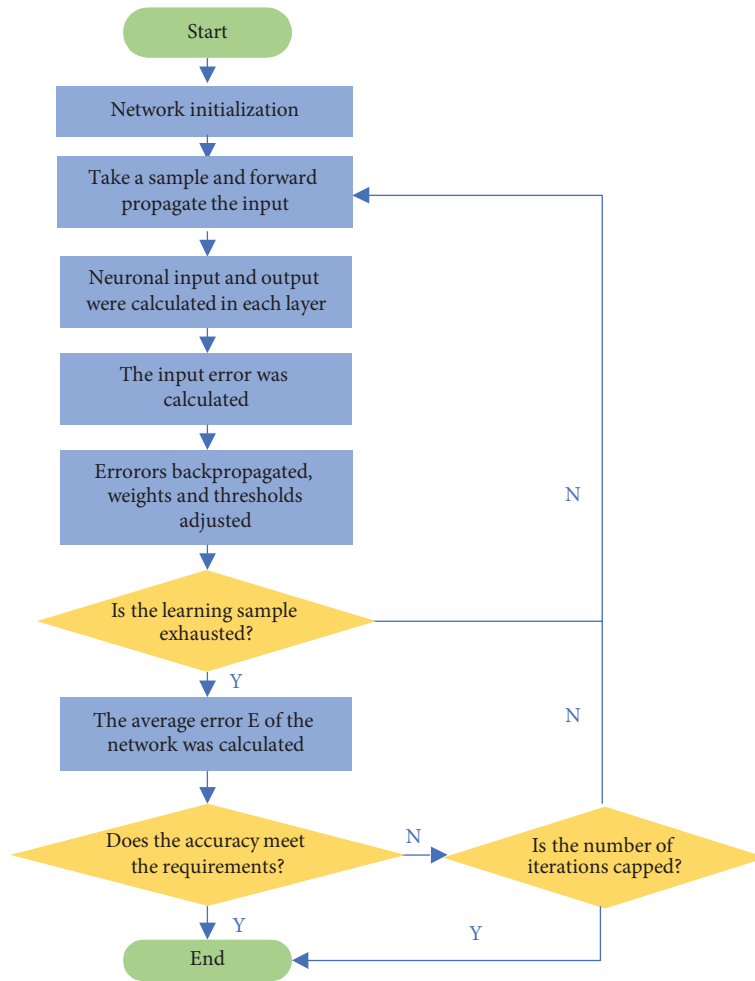
FIGURE 7: The flow of the BP algorithm.

Therefore, analyzing the driver's perception ability of risk information is the premise of designing a risk early-warning system. Driving a vehicle is an acquired skill that grows with driving experience. An ideal risk warning system should be able to reach or be better than a skilled driver's ability to perceive risks. Otherwise, if the system cannot perform better than the driver, it will lose the meaning of assisted driving. Studies have shown a big difference in driving stability between experienced and novice drivers. This experiment will quantify the driver's perception of risk to provide an effective risk warning threshold for the risk warning system. The architecture of the early-warning system based on radar sensing technology at this stage is shown in Figure 9.

Figure 9 shows the composition and architecture of the vehicle-mounted early-warning system. Laser sensors and other sensing devices collect information such as speed, acceleration, and position and upload them to the central processor for data integration and processing. The system calculates the alarm distance according to the designed early-warning model. Then, the central processing unit sends instructions to the signal output device to promptly and effectively remind the driver through sound and light on the display screen, alarm, and warning lights.
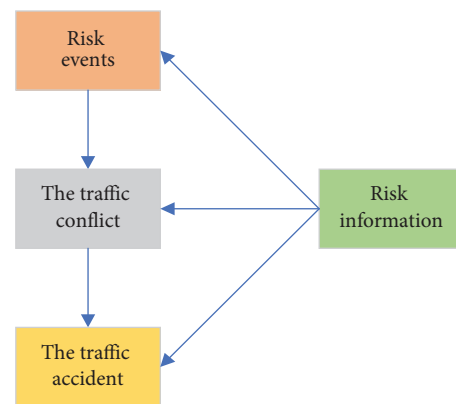
FIGURE 8: Relationship of traffic events.

6. Results and Discussion

6.1. Analysis of the Output of the Model. When a simulator drives a car, the speed display is one indicator of the accuracy of the simulation. Different places have different requirements in the driving speed test of the car. The car's speed should be 30–60 km/h when driving in an urban
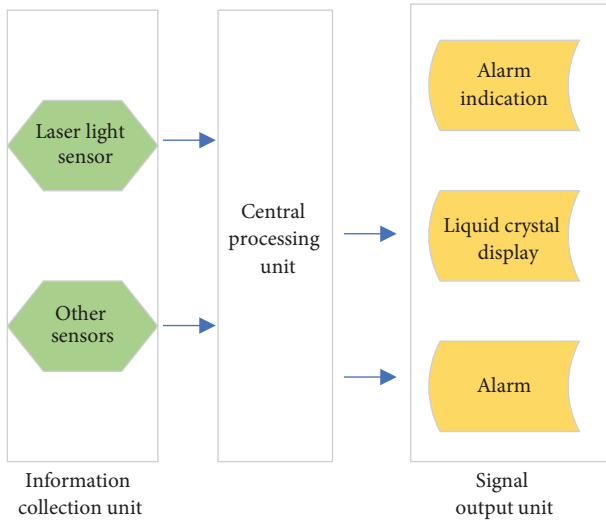
FIGURE 9: Radar-based vehicle conflict warning system.
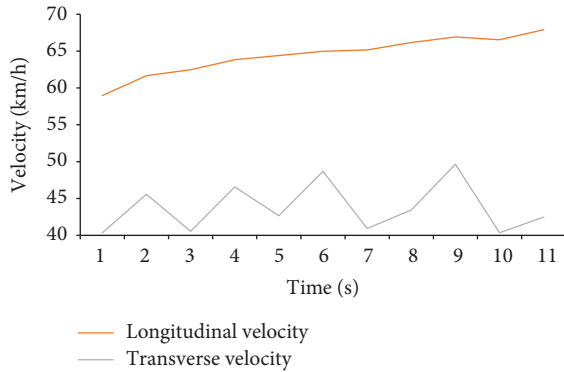
FIGURE 10: Driving speed under acceleration.

FIGURE 11: Driving acceleration under acceleration.
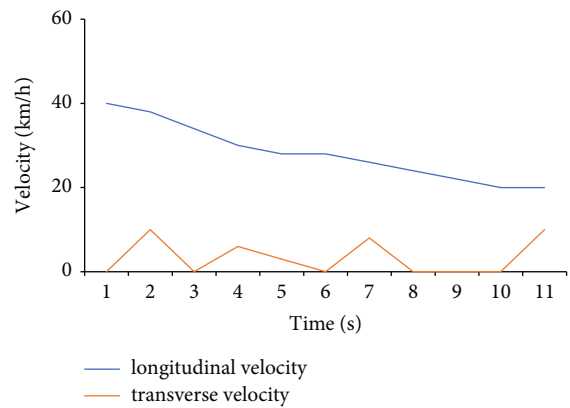
FIGURE 12: Driving speed under braking.
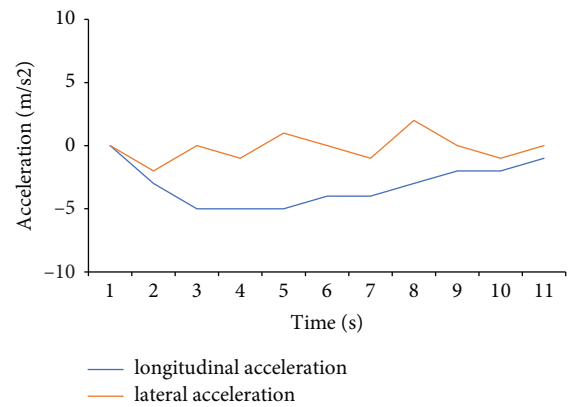
FIGURE 13: Driving acceleration under braking.

area. The traffic flow in the urban area is large, and there are many people, so the speed should not be too fast, generally at a constant speed of 40 km. The road surface of the expressway is relatively smooth and belongs to the fast lane, and the speed is generally 80–120 km/h. The road arrangement in this experiment is relatively remote for safety reasons; the traffic flow is small; and the road is smooth, so it is a relatively normal speed to maintain the vehicle speed at 60–80 km/h. Figures 10–13 show the driver's speed change based on the MATLAB simulation driving environment:

Figure 10 shows the score fluctuation of the longitudinal and horizontal speed of the simulated car under the acceleration state. The longitudinal speed has been in a state of steady growth, and the speed value is between 60 km/h and 70 km/h, which is in line with the car's speed in the actual driving process. The lateral speed fluctuates greatly, the speed is controlled below 50 km/h, and the numerical instability factor is too strong. The simulation is in good agreement with the actual situation.

Figure 11 compares the longitudinal and lateral accelerations of the simulated car under the acceleration state.

The longitudinal acceleration fluctuates less than the lateral acceleration, and the acceleration value fluctuates up and down at $5 \, m/s^2$ and is always higher than the lateral acceleration value. The lateral acceleration value is negative, and the value is between -5 and $0 \, m/s^2$. Additionally, its instability is also high. The state displayed by the data is more in line with the actual driving situation of the car.

Figure 12 is a comparison of the driving speed of the simulated car under braking. The longitudinal speed has been in a state of steady deceleration, and the value is
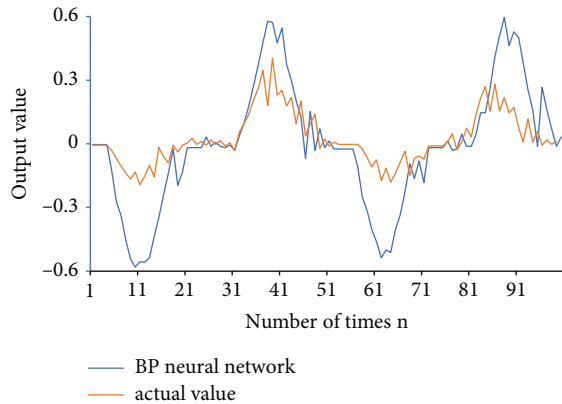
FIGURE 14: Comparison of output results.

controlled between 20 km/h and 40 km/h. The lateral speed fluctuates up and down. The situation is like the driving speed in the acceleration state. It fluctuates erratically below 20 km/h, and the lateral speed value is always lower than the longitudinal speed value. The state displayed by the data is more in line with the actual driving situation of the car.

Figure 13 compares the acceleration of a simulated car under braking. The value of longitudinal acceleration varies less than that of lateral acceleration, and the deceleration value is controlled from −5 to $0\,m/s^2$. When braking, the deceleration value is always lower than the lateral acceleration value, and the situation is relatively stable. The lateral deceleration value fluctuates around $0\,m/s^2$, and the instability is strong, which is more in line with the actual driving situation.

Among the driver's speed change values obtained by MATLAB's simulation of the driving environment, the simulation results are like the actual driving conditions of the car and have a high degree of coincidence. In the parameter value debugging process, the stability of the model is relatively good, which meets the simulation requirements. Additionally, in terms of data processing, there are relatively few comparisons of models. The car's driving conditions under acceleration and braking are compared and analyzed. The application of MATLAB in the computer simulation driver needs more practical analysis to ensure that it can be widely used in driving safety risk analysis.

7. Risk Analysis of Driving Safety Based on ANN

This paper uses an ANN model to train driving behavior, resulting in multiple individual networks. Then, the neural network dynamic integration algorithm is used for integration to establish a driving behavior model that is closer to the actual situation. The output result is the maximum weighted average of each corresponding neuron in the individual network alliance, which is the output of driving behavior. The error values of the output and the actual situation are compared. The accuracy and optimization scheme of the neural network algorithm is derived. The data results are shown in Figure 14.

Figure 14 is the output comparison between the BPNN simulation result and the actual value. The actual operating

values of the driver are simulated. After the input signal value in the network undergoes algorithm operation and program transformation, the neural network model obtains a series of data. Compared with the actual value, the error between the two is analyzed.

The error range between the BPNN simulation results and the actual value is about 0.4, and the error value is relatively low. The data show that ANN can predict and track data to a certain extent in the process of driving safety risk assessment and effectively reduce risks in actual driving safety.

The ANN dynamic ensemble learning method is used to learn driving preferences from historical driving behavior records and save them to the data system to solve the nonlinear relationship between the driving environment and driver behavior. The driving behavior is pretested in terms of the opening degree and time of the driver stepping on the brake pedal and the accelerator pedal, and the steering wheel angle and time. A possible risk indication is drawn from the historical database, and an early warning is issued. The results are consistent with the overall trend of the sample data and reflect the personality of driving behavior. Therefore, when the ANN algorithm is actually applied to driving safety risk analysis, the driver's historical driving behavior should be recorded in the early stage and stored in the database. In this way, after the corresponding situation occurs, more accurate warning prompts can be issued to ensure drivers' safe driving and road traffic.

The research is limited to the theoretical and simulation level and lacks actual data support. There are few mature theories in domestic research in this area, so the established model is still far from the actual situation. The experimental simulation data in the driving behavior research come from the driving simulator, not the data collected on the spot, and the situation of the early-warning system in the actual application process lacks verification. Therefore, the next step is to get out of the simulator, try to collect real driving behavior and driving environment data, and establish a driving behavior and a road condition database for driving behavior learning, so that this research has high practical value.

8. Conclusions

In road driving, the driver operating system is multichannel and nonlinear. The MATLAB driving simulator is used to analyze the car's dynamic characteristics during driving and the possible factors of risk occurrence. Combined with the simulated driving environment and the speed change value of the driver, the simulation results are like the actual driving conditions of the car, with a high degree of coincidence. In the parameter value debugging process, the stability of the model is relatively good, which meets the simulation requirements.

The application of ANN in driving safety risk warning is combined with an example of risk prediction. Deep learning, combined with the actual driving situation, analyzes the possibility that the car may have a risk in some cases and issues an early warning based on the ANN operating mechanism to reduce the possibility and injury of traffic accidents. The BPNN is used to model the behavioral

characteristics of drivers. The neural network system is practiced and trained many times. The test results are compared with the driver's actual risk value. The error range between the BPNN simulation results and the actual value is about 0.4, and the error value is relatively low. The BPNN model can track the changes of different instructions very well. For imminent hazards, the model warns the driver and exhibits certain predictive and data-tracking capabilities.

Additionally, the data processor also showed certain shortcomings. There are relatively few comparisons of the models. This work only compares and analyzes the driving situation of the car under acceleration and braking state. In the later stage, MATLAB needs to do more practical analysis on the application of computer simulation drivers to ensure that it can be widely used in driving safety risk analysis.

Data Availability

All data used to support the findings of the study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Witt, K. Kompaß, L. Wang, R. Kates, M. Mai, and G. Prokop, "Driver profiling–data-based identification of driver behavior dimensions and affecting driver characteristics for multi-agent traffic simulation," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 64, pp. 361–376, 2019.

[2] J. W. C. Van Lint and S. C. Calvert, "A generic multi-level framework for microscopic traffic simulation—theory and an example case in modelling driver distraction," *Transportation Research Part B: Methodological*, vol. 117, pp. 63–86, 2018.

[3] V. Astarita and V. P. Giofré, "From traffic conflict simulation to traffic crash simulation: introducing traffic safety indicators based on the explicit simulation of potential driver errors," *Simulation Modelling Practice and Theory*, vol. 94, pp. 215–236, 2019.

[4] Q. Luo, X. Chen, J. Yuan, X. Zang, J. Yang, and J. Chen, "Study and simulation analysis of vehicle rear-end collision model considering dri ver types," *Journal of Advanced Transportation*, vol. 2020, pp. 1–11, 2020.

[5] J. Sini and M. Violante, "A simulation-based methodology for aiding advanced driver assistance systems hazard analysis and risk assessment," *Microelectronics Reliability*, vol. 109, Article ID 113661, 2020.

[6] S. Amin, "Back-propagation–Artificial neural network (BP-ANN): understanding gender characteristics of older driver accidents in west midlands of United Kingdom," *Safety Science*, vol. 122, Article ID 104539, 2020.

[7] M. Watanabe and K. Sakai, "Identifying tractor overturning scenarios using a driving simulator with a motion system," *Biosystems Engineering*, vol. 210, pp. 261–270, 2021.

[8] M. AbdElSalam, K. Khalil, J. Stickley, A. Salem, and B. Loye, "Verification of advanced driver assistance systems (adas) and autonomous vehicles with hardware emulation-in-the-loop A case study with multiple ECUs," *International journal of automotive engineering*, vol. 10, no. 2, pp. 20194104–20194204, 2019.

[9] A. Depari, E. Sisinni, A. Flammini et al., "Autobalancing analog front end for full-range differential capacitive sensing," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 4, pp. 885–893, 2018.

[10] J. S. Wijnands, J. Thompson, G. D. P. A. Aschwanden, and M. Stevenson, "Identifying behavioural change among drivers using Long Short-Term Memory recurrent neural networks," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 53, pp. 34–49, 2018.

[11] X. Rao, F. Lin, Z. Chen, and J. Zhao, "Distracted driving recognition method based on deep convolutional neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 193–200, 2021.

[12] A. M. Amiri, A. Sadri, N. Nadimi, and M. Shams, "A comparison between artificial neural network and hybrid intelligent genetic algorithm in predicting the severity of fixed object crashes among elderly drivers," *Accident Analysis & Prevention*, vol. 138, Article ID 105468, 2020.

[13] J. Yuan, M. Abdel-Aty, Y. Gong, and Q. Cai, "Real-time crash risk prediction using long short-term memory recurrent neural network," *Transportation Research Record*, vol. 2673, no. 4, pp. 314–326, 2019.

[14] Y. Huang and S. Meng, "Automobile insurance classification ratemaking based on telematics driving data," *Decision Support Systems*, vol. 127, Article ID 113156, 2019.

[15] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural networks," *Expert Systems with Applications*, vol. 149, Article ID 113240, 2020.

[16] Q. Shangguan, T. Fu, J. Wang, T. Luo, and S. Fang, "An integrated methodology for real-time driving risk status prediction using naturalist ic driving data," *Accident Analysis & Prevention*, vol. 156, Article ID 106122, 2021.

[17] J. S. Wijnands, J. Thompson, K. A. Nice, G. D. P. A. Aschwanden, and M. Stevenson, "Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks," *Neural Computing & Applications*, vol. 32, no. 13, pp. 9731–9743, 2020.

[18] 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.

[19] W. Zhao, L. Xu, J. Bai, M. Ji, and T. Runge, "Sensor-based risk perception ability network design for drivers in snow and ice environmental freeway: a deep learning and rough sets approach," *Soft Computing*, vol. 22, no. 5, pp. 1457–1466, 2018.

[20] J. Peng and Y. Shao, "Intelligent method for identifying driving risk based on V2V multisource big data," *Complexity*, vol. 2018, pp. 1–9, 2018.

[21] R. Ahmadi and S. A. Hamdan Shybt, "Study of artificial neural networks in information security risk assessment," *Journal of Management and Accounting Studies*, vol. 8, no. 2, pp. 1–10, 2020.

[22] J. Lee, J. Kim, I. Kim, and K. Han, "Cyber threat detection based on artificial neural networks using event profiles," *IEEE Access*, vol. 7, pp. 165607–165626, 2019.

[23] J. Pacheco, V. H. Benitez, L. C. Felix-Herran, and P. Satam, "Artificial neural networks-based intrusion detection system for internet of things fog nodes," *IEEE Access*, vol. 8, pp. 73907–73918, 2020.

[24] L. Y. Leong, T. S. Hew, K. B. Ooi, and J. Wei, "Predicting mobile wallet resistance: a two-staged structural equation modeling-artificial neural network approach," *International Journal of Information Management*, vol. 51, Article ID 102047, 2020.

[25] I. Sumaiya Thaseen, J. Saira Banu, K. Lavanya, M. Rukunuddin Ghalib, and K. Abhishek, "An integrated intrusion detection system using correlation-based attribute selection and artificial neural network," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 2, p. 4014, 2021.

Hindawi

Research Article

Enhance Software-Defined Network Security with IoT for Strengthen the Encryption of Information Access Control

Vrince Vimal,[1] R. Muruganantham,[2] R. Prabha,[3] A. N. Arularasan,[4] P. Nandal,[5] K. Chanthirasekaran,[6] and Gopi Reddy Ranabothu [ID][7]

[1]Department of Computer Science & Engineering, Graphic Era Hill University, Dehradun, Uttarakhand, India
[2]Department of Information Technology, TKR College of Engineering & Technology, Meerpet, Telangana, India
[3]Department of ECE, Sri Sai Ram Institute of Technology, Tambaram West, Chennai, Tamil Nadu 600044, India
[4]Department of Artificial Intelligence and Data Science, Panimalar Engineering College, Chennai, India
[5]Department of Computer Science and Engineering, Maharaja Surajamal Institute of Technology, New Delhi, India
[6]Department of ECE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India
[7]Wachemo University, Hosanna, Ethiopia

Correspondence should be addressed to Gopi Reddy Ranabothu; gopir@wcu.edu.et

The Internet of Things (IoT) is legitimately growing quicker. The operators have already started setting up a diligent infrastructure for these gadgets. Various technologies need to be developed for this type of sensor, including enterprise safety initiatives. This paper covers the stability routing protocol, which assumes an assessment of credibility in gadgets and packet flow. To build reliable Software-Defined Network (SDN) routes, build on the trust between network element flows and Quality of Service (QoS) or energy conditions. The SDN architecture is used for the Cognitive Protocol Network (CPN) technical platform to increase the energy level. Stochastic Neural Networks (SNNs) are accredited with information extracted from perceptual packets and make decentralized decisions. The proposed network infrastructure is designed and integrated into the SerIoT techniques to strengthen IoT encryption for information access control. The versatility of the technology is to circumvent the unpredictable connectivity of the system and the node decreases in terms of potential cryptographic capacity, limited interval, a target node, and deterministic energy. Based on factual statistical data, appropriate marketing generates an end-to-end antitheft solution that meets a set of predetermined circuit restrictions. A study must collaborate by demonstrating numerous flaws due to the obvious instability of clusters, which is essential for the efficiency of the platform.

1. Introduction

Stability in IoT is no longer perceived as a secondary concern, but rather as a relevant concept throughout the development of technological infrastructure or software applications. Intimate location, biometric data, or business information are among the most dangerous data acquired by electronic objects. Interconnections extract data that would be used to control urban and industrial infrastructure [1]. The deteriorating IoT sector raises concerns about the reliability of the IoT devices or the connections that send information. If

verified or corrected, then aggressions have an enormous cost to the distribution system, including the lack of credibility in a competitive world [2]. Manage devices that have hacked abandonment risk based on customer or end-user credibility to increase overhead, legal costs, increased electricity usage, operating costs, and CO_2 damage [3].

Partitioning is one of the most important steps to ensure secure Internet access. As a result, hackers focus on targeting 14,000 attacks such as drains in data loss, income loss, and reputational damage in 2017. Navigation assaults include a malware on the delivery service, intrusion on the

morphological provider, assaults on several path services, and facilitation of authenticity. It was categorized in [4, 5] and intends to exhaust network capacity and espionage. Track protection requires considerable investigation. Configuration as a technique of assault prevention gets very little attention in the search for the difficulty that arises in ad hoc networks. A new website on protection-sensitive sequencing in SDN base stations [2] is already disclosed.

Traffic Analysis and Fault Detection (TA/FD) are surveillance or mitigation strategies commonly used in computer systems. It stipulates the adequate encryption methodology used to protect the most vulnerable components of the network where TA/FD is not able to make a definitive remediation option [6]. Leads potentially vulnerable parts to the transportation of dangerous goods system to provide recreation for thorough investigation or redirect flows [7].

Rerouting the importance of navigation in preventing assaults in computer systems led us to include the administration of security circuits. As a key component of SerIoT (Reliable or Protect Wearable computing) [8]. Collaborates on unique channel management approaches for SDN networks that focus primarily on the monitoring and disclosure of online perceptual security [9].

Accumulate numerous safety statistics or have clues about restricted insurance of specific vehicles or hubs that redirect pedestrians presumed to be components of a cyber incident to a less immediate alert [10]. Adapting protective trails or omitting the server are all examples of trail monitoring based on stability. To achieve improved durability assaults as a response to newly established techniques or methodologies, raising a measure of assurance for either network provider or network consumer gadgets [11]. One reason for focusing on SDN's substantial progress in investigation or installation, penetration testers are still in their infancy for comparatively small and talented innovation.

More specifically, the EU's FP7 NEMESYS research allowed researchers to study the stability of mobile operators. It includes the centralized controller component of the functioning of the mobile network, and several attacks specifically target it [12, 13]. Vulnerabilities in infrastructure technologies that exploit the research when it is functioning correctly or in danger. Messaging services used to leverage data, information, and instructions from various parts of the network are all presumed in the further investigation of surveillance of malware processes [14].

In [15], a relevant current study on cybercrime in Europe is reviewed on numerous initiatives funded by the European Commission. Cybercrime for mobile communications is a serious barrier to all aspects of information systems [16]. While the majority of modern mobile devices offer WiFi access from other cellular routers, the emerging security controls should be continuously monitored on the network and dominate aircraft levels in smartphones. Consequently, a previous survey [17] examined the use of Artificial Neural Networks (ANNs) and computer vision technologies to address this concern. Research [18] focuses on assaults on the aircraft triggering the foundation network, which has a direct impact on the mobile service provider or the end consumer. The NEMESYS initiative has dispelled many of these difficulties by leveraging the methodologies of queuing theory.

Concentrates [19] on the integrity of communications and information delivery for European regional or national medical systems which are equipped. European explorers occasionally need to seek medical coverage in another European country, health computing platforms should be able to get enough virtual patient records. Stability in the residential IoT industry is addressed in [20], which focuses on the architecture of a robust residential central server. It includes assault identification techniques and evaluation of assault tactics that attempt to deplete energy production from products by exhausting their devices. The SerIoT initiative began in January 2018, and further information may be found [21]. To persuade or investigate attacks, it should develop and execute phishing emails on virtual servers.

2. Materials and Methods

SernCPN, a CPN-based network that affects custom dispersed IoT systems, would be discovered through research. This would use calibration to create system self-awareness by implementing the SDN built on the CPN. This SDN would use Cognitive Packet (CP) to find secure multihop paths with a reliable Internet connection. And, evaluate their security and reliability, supervised learning with SNN to improve the overall performance of the system. It would include the three objectives of safety measures, good service, and lower energy consumption. Multiple sets of SernCPN networks can be credited together using a final integration node and adaptable links to Grid or Cloud platforms for network data management and analytics.

In addition, AI and IoT could be leveraged in coordination to remove the ZeroDay assault. A zero-day exploit exposes an organization to previously unexplored vulnerabilities, and there is no way to build a counterforce. A zero-day exploit exposes an organization to previously uncharted vulnerabilities, and there is no way to build a counterforce. Artificial intelligence (AI) would be designed to monitor the rate of potential malware with features derived from relevant services that predict the probability of traps. Meanwhile, AI and IoT could be used synergistically to impede Advanced Persistent Threats (APTs) based on adaptation classification methods [22]. In one study [23], one of the cases involving the integration of IoT and AI for enhanced security is explored. The authors of this research guideline an AI approach to reduce banking crime in an IoT scenario. This article highlights how IoT leverages recruitment and evolution algorithms to determine customer behavior in securities fraud prevention, and credit card theft, using an approach like BOAT [24]. According to Choi and Lee's research, using an ANN could accurately detect a forgery in actual time.

3. Design Theory of Secure IOT System

An IoT system's creative process is especially important in industrial platforms since it dictates which equipment a company would navigate to cognitive computing [25]. These

gadgets, in turn, are part of the enormous particles of data surveillance of the universe. The same reliability of enterprise-critical data would be dictated by the installation of IoT systems. IoT's creative process would have implications for safety protocols because it depends on a combination of offspring. Include a well-articulated communications plan, reinvent accountabilities when developing a security protocol plan, and keep them informed of key innovations to ensure information security [26]. In addition, Protected IoT's research approach is based on three primary criteria. The first step involves connecting direct input devices or embedded annotation devices to the main cloud server. This includes multi-factor authentication of responders and verifying equipment attempting to enter the facility. A second step would allow retrieving, optimizing, and delivering substantially the data of different instruments in the gadgets. This is achieved by choosing the optimal data or model that is acceptable to the cloud-based industrial IoT architecture.

One of the goals of the research is to build and test a reliable SDN-related IoT fiber network and an intelligent controller with Internet perceptual monitoring and disclosure. The ability to build and revise routes in an adaptive way to increase security for IoT devices and business customers while preserving the near-optimal quality of goods. CPN methodology introduced in [27], is used for online perceptual supervision and route surveillance. Numerous studies, discuss the principal components or functionality of the CPN router. The SernIoT "SernIoT CPN network" or SernCPN smart SDN network starts with some of the concepts described in [28].

The study of a revolutionary approach to transmission involves related concepts. Although the term may seem ambiguous in practice, many researchers have already defined confidence in a form that is also adopted for the study. A "sustainability intensity" of nodes within a network is considered to be trusted [20]. According to this description, credibility is defined as the likelihood that collaborating nodes in a network will adhere to the network's regulatory frameworks. It would not violate the security criteria of anonymity, stability, access, legitimacy, and nonrepudiation.

SernCPN would be a secure multipurpose network platform, concerned about the QoS or energy that could be used in various scenarios, including the IoT domain [20].

(1) A digital IoT platform would be separated from the contractor's base network

(2) Overlay on the Internet, in which regular Internet capabilities are used instead of wired connections

(3) Local communication through a major IoT network

SernCPN takes the baseline SDN method and enhances it. It features data and configuration control plans and interfaces between them using the OpenFlow protocol. Priority-based eligibility guidelines should be used to make decisions about data flow:

(1) Security & Stability: information must be provided securely, with the least impact of being blocked and lost (due to deliberate or accidental defeat). This includes safeguards against hacking the switch and regulator and measures to pass erroneous data to a central device (or other networking devices).

(2) QoS: manifestations such as flow, latency, and fluctuation are important criteria in determining packet arrival routes, and this could be done using CPN.Fused Goal Functions (FCF) that contain encryption, QoS, and possibly energy, as illustrated below.SernCPN should be secure, but its QoS should be attractive to customers.

(3) Energy utilization: when choosing particle trajectories, efficiency would be considered. A charge on the switches will be made to reduce power consumption, and traffic will be scattered on roads to reduce power consumption per packet or link [29].

The navigational provisions are reflected in an SDN-based SernCPN acquisition by determining the appropriate rules for the specified loads. An "oracle" would make the relevant decision, which would be powered with protection, QoS, energetic data, and stored in an Intellectual Security Memory (ISM). It would be built using Recurrent Neural Networks (RNNs) using a major learning technique such as Evolutionary algorithms in CPNs. A handler plugin would contain RNNs and ISM data used in machine learning. It should be noted that the RNN Reinforcement Learning (RL) based formative evaluation not only leverages current sensor readings.

Under the conditions described in II-B, ISM data will be classified into three parts. A great majority of user attributes, their definition for the evaluation process of destination paths and IP addresses, would be developed throughout the quality of the search.

It should concentrate on the statistical measures or assessments used by RNNs for recycling and then forward them to the SerIoT monitoring elements for monetization. Many methods or methodologies exist for traffic forecasting and detection systems [30]. Advanced threat segmentation methods would be selected at a later stage of the research. But it should begin with a range of topics that would give us a statistical perspective on the security connectivity of network devices.

(1) Assurance assessment (reliability testing) of systems connected to a SerIoT device: the probability of being the source or recipient of the attack. Reasonable precautions are considered in this category. Basic confirmations could involve the design of default passwords, and profiling firmware versions about the most modern security updates deployed, but more complicated.

(2) SFE Security (Trusted Level): a node is intended to be hacked, disabled, or monitored. It could take advantage of the level of accessibility of servers as a strategy to adjust the overall credibility of a circuit.

(3) Security for specific purposes is achieved through a variety of methods. The detection or use of assays, for

example, is a basic indicator (lighter weight) that a given flow could be a component of assault.

(i) Verify if the source or destination IP addresses of the stream are listed in a public list of forbidden IP addresses to be targets of aggression

(ii) Identifying flows over a predetermined limit

(iii) Access the IP address using the acquisition

(iv) Identify the plan for using nonstandard connections to apply effective techniques, such as comparing data traffic with attack methods using computational means or cognitive computation, at the end of the research

(4) QoS attributes are assigned through specific channels. The performance of specific lines, latency, instability, and loss levels should all be monitored by CPN.

(5) The energy consumption of certain nodes in terms of actual traffic volumes is determined.

Consequently, the FGF for SernCPN optimization would be created using the elements indicated in (1) to (5).

Intellectual Goal-Setting $G(f, P)$ accepts nonnegative numerical actual values, where f symbolizes a flow (traffic defined by its source node, leading to impaired, broadcast address, and destination) and P represents a particular network route. Therefore, the quantities have been determined by the flow to which it relates the router cost. In either case, the decision system should search for a new flow path, leading to a lower G value. The entire function will also take QoS and effectiveness into consideration. It also describes responsibly the situations in which it is located.

The rejection level of flow f as in-circuit e R is now determined (f, e). The rejection factor is calculated using equations (1)–(3).

$$S(e, f) = \begin{cases} 0 & \text{If } RF(e, f) \le SE(e, f) \\ \text{Rejection Factor}(e, f) - RE(e, f) \\ \text{Rejection Factor}(e, f) > RE(e, f), \end{cases} \quad (1)$$

$$\begin{aligned} & \text{Rejection Factor}(e, Q) \\ & = \sum_Q S(e, f), \text{ or } S(e, Q) = \text{maximum of } S(e, f), \end{aligned} \quad (2)$$

$$H(e, Q) = \text{Rejection factor}(e, Q) + P(e, Q) + F(e, Q), \quad (3)$$

where $S(e, f)$ represents circuit at reception level e of flow f; Q represents path; P and F are non-negative components.

Private Multihoming Protocol: as our reliance on digital grows, and the complexity of existing systems grows, it becomes increasingly difficult for a gadget to work alone. Cluster-driven interactions are required in this situation, which could be achieved via a range of approaches, including compound broadcast and multiplex. Each ensemble node must receive a point-to-point parcel when using compound multicast. Multipathing is a theoretically complex function

that allows you to send a package to multiple devices while safeguarding your privacy by disseminating the message only as far as it has to go to impact each ensemble node, and only once on each path in Figure 1. To study shielded multiplex systems, numerous standards are developed, including ensemble affiliation maintenance; connect capacity absorption, receiver material demands, respondent workforce demands, and dependence on particular criteria.

If communications are required for an all-inclusive ensemble as well as for portions of a population, IIoT allowed safe multipath routing erupts. Protected multipathing is operationally better when the constellation would be sufficiently stationary due to the overhead (in Polynomial-based Key Management) of a larger number of rekeying routers every moment a node links or leaves that has a higher registration level than the lowest point as in cascade. They assume that multicast groups are always built using the core-based tree technique, in which each multicast pattern node has an unrevealed key that is shared with the central hub. It is worth noting that a central node is in charge of key distribution. A requirement for nodes to adopt multiple keys could express concern about node capabilities. Some nodes in an assorted ensemble have more attention resources than others. As a result, the requestor would first endorse a cluster controller before connecting to an ensemble to get any secrets. A validating router could act as a mediator, enabling new transactions to take place. Table 1 highlights the elements of necessity and preliminary Figure 2.

4. System Architecture

Routing protocol techniques would be used to enhance a traditional SDN network. They would offer new SernCPN elements that are capable of performing activities required to achieve a research's objectives, especially intelligent safety-conscious forwarding. SernCPN's major parts are: SFE switches packages frequently by Open Flow regulations. SFE would also leverage the CPNs methodology to collect data on safety, quality of service, and energy demand.

SernCPN used a standard SDN controller in combination with SernCPN Navigation Engine (SRE). The RNN-based Intellectual Routing Module at the basis of SRE makes routing protocol using the methodology mentioned.

SernCPN combined SernCPN Routing Engine with a conventional SDN controller (ONOS–https://onosresearch.org/) (SRE). SRE's RNN-based Conceptual Routing Module creates congestion control using the methods presented.

SernIoT Honey trap is another platform that emulates the suitability of different products; it is linked to SernCPN and assesses the threats it receives. It could be taken by an adversary while causing damage to other SernCPN units.

In both IIoT setups or the communications system for transferring data to industrial integrated systems, a router is inserted into a circuit. It consists of a perturbation criterion, that would be used to train a system workplace to perform intrusion revelation on layer 3 (Internet layer, which could also endorse Multicast Broadband Service protocol or Network Technology) and layer 4 (information connection, that could endorse Multipath TCP) of a modus credo field,
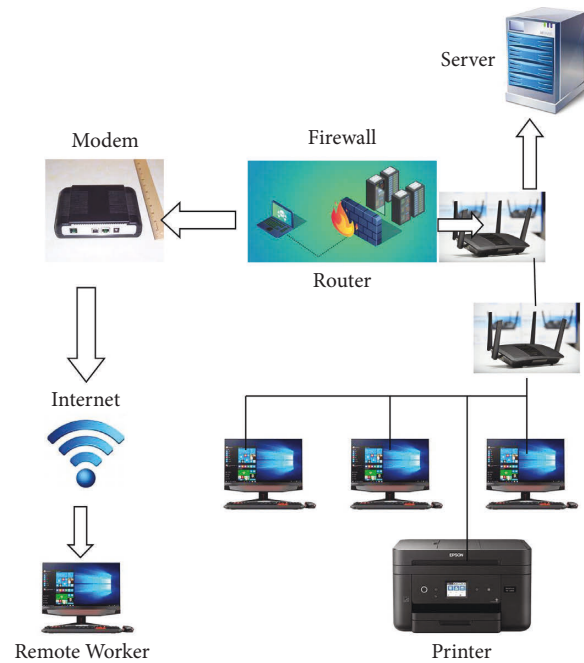
FIGURE 1: Sample diagram of basic single domain network.

among data in a routing form of packets conveyed to web. It is worth mentioning that pathway description aggregates a network port's contents, which would be segregated into a series of packets containing the initial required information and customized to a chosen messaging protocol.

Unimaginable Particle Analysis enables network monitoring, controller provisioning, and security features, as well as constellation data gathering for outlier detection. A United States Patent and Trademark Act (UPA) makes it possible to determine who originated or received a data fingerprint containing specific sequences. Moreover, UPA could disclose node efficiency, incinerate network behavior, or support service providers in augmenting throughput or outcome. Packet processing rules and characteristics are provided in Tables 2 and 3.

Rigorous match: this type of resemblance involves exact comparability of detector element, i.e., a detector sector has one value provided in it.

Range match: an examination of a packet header to be in an array regulated by the classifier is known as a speculation match.

Protocol neutrality: a methodology that passes through a purifier should be independent, so that processing could be adjusted for different procedures or phases.

Protocol Fragmentation: a packet filter could handle packet decomposition. Effective regulation informs or audits: A design allows for integration or elimination of operations with the least amount of destruction to packet delivery. It keeps track of all entrance attempts, both hypothetical or blocked if required, as well as data that could be useful for study.

Considerations ranking: if a data set meets many syllabuses, packet criteria would allow objective relevance to be enforced on graded regulations, resulting in a distinct policy being in premise essential.

SE process entails: (a) inspecting data to determine whether each of a predetermined set attributes of maliciousness exists; (b) determining a groove focused on the existence or lack of a set of rules in data, with merit/ratio delving deeper into the likelihood that data is abusive; or (c) dispensing to record based on data ratio. Support Vector Machine (SVM) is used in the suggested method, which would a popular technique for information categorization or advancement that has also been successfully used in malware detection. The SVM generates nonlinear restrictions in the flesh database by creating an evenhanded state border in euclidean space. Figure 3 depicts a process flow diagram.

5. Implementation

SernCPN Routing Engine (SRE) would be redistributed as a plugin element in one or more SDN controllers, leveraging RNNs to build judgment oracles, allowing a semidistributed way of making judgments while leveraging the benefits of the SDN platform's semicentralization. A physical network topology would be replicated by attaching certain RNNs to SFEs. A single RNN's job would be to indicate, at the moment of decision, that the outlet unit should be used for a given SFE in the context of a flow with a particular destination. SRE would collect information in two aspects:

(1) Clever Packets are an example of this (SPs)

(2) through a supervisor that collects data on surveillance or diagnostic units

SerIoT, which would be based on the CPN concept, employs Intellectual Packets that go from one node to another or on their way to finding and collecting recorded network nodes. Typically, the SFEs of nodes accessed by CPs supply a CPs' path, while ACKs' path is origin rerouted from the destination address. The data generated by each CP that

TABLE 1: Parameters of necessity.

Title	Unit	Scheme assumed
Operational simplicity	Set the timer interval	Clusters of compound compounds
Intensity communications in real-time	Buffer for addresses	Enable
Depletion of resources	Distributing access	There are several entries for each link.
Per message connection rate	A network that is only loosely connected	Extreme
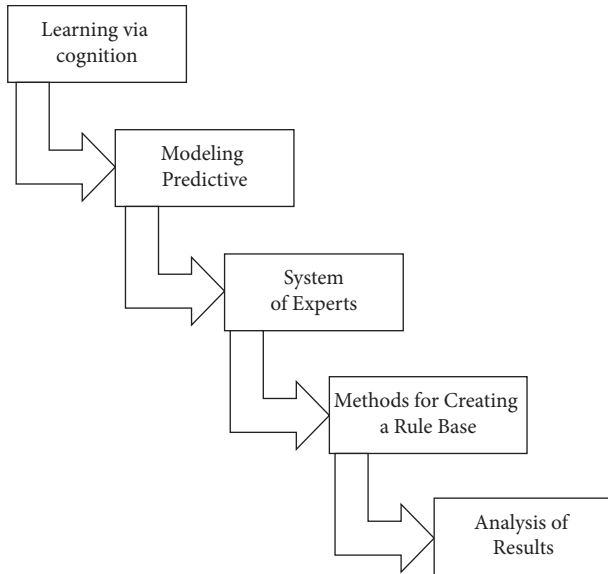
FIGURE 2: Steps to implementing context-aware computing.

TABLE 2: Environmental parameters filtering.

Parameters	Units
Size per packets	36 bits
Route time	2000 ms
Simulated kit	ESP8266
Purpose	12
Latency	120 ns
Input	300 megabytes per second
Packet filter for nonlinear	50 megahertz
Destination port	500

accessed the node could store or exploited by such hubs. A technique is amended in SernCPN since transmission of CPs or navigating throughout the network is managed by a regulator or SRE, so ACK messages, rather than flowing through the system using a path back to its source node, go to SRE that uses their information to determine navigation.

CPs would be used for information that would otherwise be unavailable, like connection delay or overall delay of neighboring nodes, including delay inside hubs, as well as data that could be provided directly by nodes (Web Sockets or demand) but is less essential (e.g., energy usage). CPs combine data to multiple sites along their journey and convey it to the supervisor in a short statement, decreasing communication costs.

Codesign would be the study of people interacting with workstations or to what extent workspaces or not urbanized to effectively communicate with humans and employees. One important element of codesign would be that different services with different outsets could replicate their cooperation and have different learning and customer attentiveness characteristics. Artificial intelligence, on the contrary, would be an imitation of general intelligence advancements using technology, primarily workstation PCs. Epistemology (gathering evidence or using syllabuses to expend data), sensation (using instructions to affect derivation or specific presumptions), or sympathetic activity are examples of these methods. Turing Test is a very well method for determining if technology could think like a person. They employed a classification algorithm to

portray document matrices such that margins could be detected and used to tag different track assemblages. It is worth noting that while developing our approach, they took into account to General Data Protection Regulation (GDPR) published by European Union (GDPR). Several trials are performed to validate a proposed IIoT-enabled robust encryption. Table 4 provides an overview of the experimental situation.

This appraisal is based on two sets of studies: (a) demonstration of recurring safety or (b) affirmation of resource-conscious safety. Table 1 shows that possibility of detecting a hazard is dependent on a platform's ability to detect an abnormality appropriately. A criteria precision, detection ratio, or deceptive level of confidence could be used to evaluate methods. The condition of payload delivered across the boarding gate establishes corresponding statistics; this categorization is expressed by labels true/false positive/negative.

During the trial, they focused on gadget surveillance, configuration encounters, connectivity, infrastructure protection, and IIoT framework based on the test best habitat. They can isolate a probability of cyber threat forecast modeling using data collected for Table 4. They discover that gadget polymorphism might mimic the overall benefit of converting policies into maintaining levels, as distinct categories of assets might well be beneficial to combat associated dangers. Various methods/policies permit an indication of unique dissimilar peculiarities of the process to minimize false warnings depending on the exact quality requirements under monitoring or existing facts. Regularly enhance system conditions depending on efficiency challenges during the trial.

For a myriad of purposes, a large number of different things attached to an IoT device creates significant security risks as shown in Figures 4–6. Imperfect experiences assess anticipated IIoT test-bed /debugging in terms of energy and information processing qualities. The chart shows an overview of simulator tests, that is developed to generate a generic threshold for security procedures in IIoT aligned

TABLE 3: Criteria for evaluating malware.

Malware scoring principles		Malware grouping rankings	
Average	Storyline	Parameters	Description
Launch	The program component is on the verge of being implemented.	Serious	Storage is disrupted.
Circulation	The cipher has been disseminated.	Hazardous	Bandwidth in distress
Level of resentment	The payload is connected to the set on.	Negligible	Easily accessible and useful
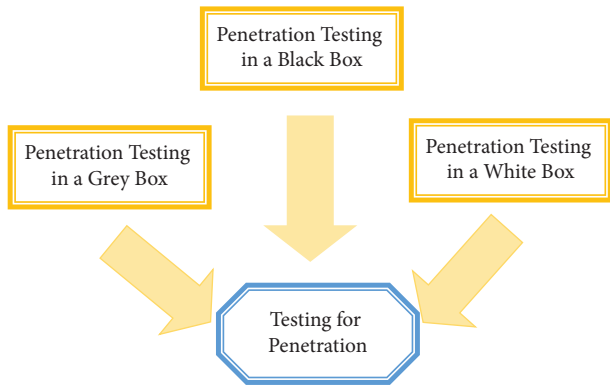
FIGURE 3: Analysis of vulnerability assessments.

TABLE 4: Environment for trial and error.

Parameters	Units
Test	Regular grid
Configurations	5
IIoT sensors	150
Node capacity	50 GB
Storage capacity	20 TB
Pins	20
SoC	Adafruit FONA
Network	802.11 bgn
Antenna model	Low power omni direction
Signal propagation	Antenna with spring
Software	Arduino IDE
Cloud	Microsoft
RAM	DDR–3, 4 GB
Processor	Quad-core
Port number	8
Payload	60 bits
Duration	150 mins
Number of events	6000 per 15 minutes
Humidity	25%
Temperature	−3 minute
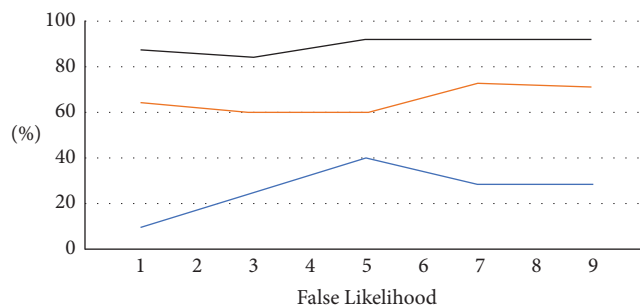
FIGURE 4: False and threats of likelihood.
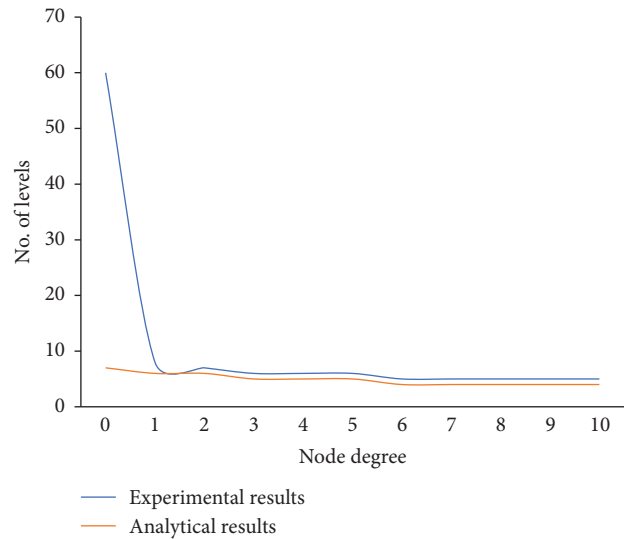
FIGURE 5: Results of experimental and analytical values for IIoT provision.

anomalous mitigation context. The main goal of this experiment is to examine at edge nodes behaved during the outlier detection learning stage on the operating system, data tier, and Internet, as well as intrusion detection systems got better. A methodology assisted homeostatic nodes in handling a range of threats in a context-aware IIoT defensive scenario by providing an attempt to learn authority on allocating security capabilities within security components in Figure 4.

An overall handling period or security requirement fulfillment ratio is evidenced in Figure 5 based on the returning degree of security criteria for a variety of values of the system call's finishing time. It shows the results of scattering to defense mechanism on warning channels in this section. It constructed an instance because detected abnormalities are not always assaults, but could be a wide range of technical issues. It exploits by squeezing out alarm problems deliberately so that system components could be used for more relevant alerts. As a result, sorting is considered just as essential as abnormality detection. A suggested technique creates a stunning visual representation of the endangered service's security level. An accompanying chart illustrates the IIoT platform's security posture in an easy-to-understand manner. The handler could make predictions about hostile activities shown by the defensive warning axis in Figure 6. If an uneasy behavior is predictable, the context processor could check the alarms policy to ensure that event is rigorous (DAM classification).

FIGURE 6: Alert processing time and success ratio.
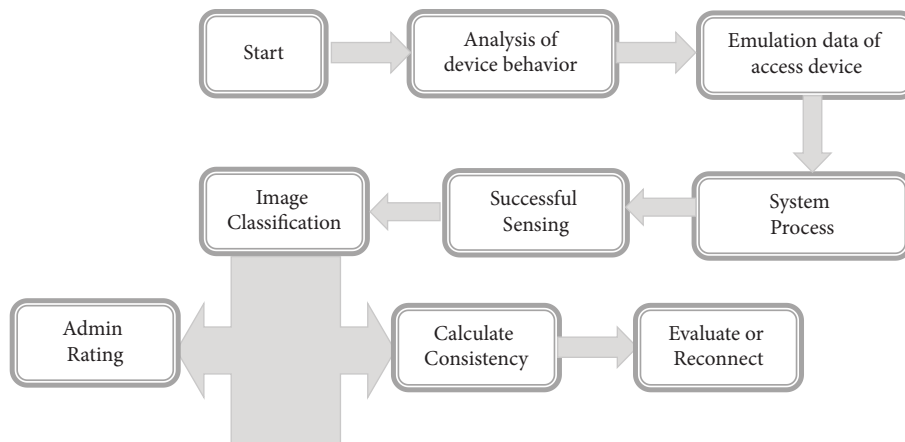
FIGURE 7: Modeling of notification trust.

The trust value notification modelled is shown in Figure 7. The time is taken to deliver data from the queue or construct a data frame not been estimated because it is extraneous to the data's range. To avoid connectivity or processing latency, set a 260 alerts/sec/zone criterion after a rigorous unit description study. This contract's receiving ratio of 9000 warnings per second is substantially higher than the gadget revealing that its technology can manage expandable IIoT networks with greater warning recurrence rates.

6. Conclusions

SernIoT is to enhance a wide range of security techniques targeted to a quickly evolving IoT ecosystem. A key emphasis of techniques given in the study is attributed to the SernIoT initiative by authors is on novel protocol methods in terms of quality and integrity. The impact of route discovery on the security of IoT devices linked to signal integrity core network equipment is evaluated. These features of network activity via IoT devices would allow us to more clearly characterize dangers and indications of malicious actions than in a general network. The proposed model is designed for security-related data across multiple risks and possible attacks. As a result, the RNN component could be used as an intelligent tool for making appropriate decisions in a dynamic environment. Protection represents an edge node and the analysis predicted remedies for mitigating security risks. The proposed system enhanced security capabilities or programs with secure or user-friendly graphical user interfaces. A potential mitigation mechanism reduces the likelihood of safety threats and incidents. In a moderate approach, effective self-governing or parasympathetic protection displayed a positive influence on overall system efficiency. The focus of this study is on identifying potential risks, impacts, or dangers, and preventive actions for IIoT systems. The goal

of our investigation purpose is to provide a framework to understand or monitor security vulnerabilities in the IIoTdefense shield.

Data Availability

The data shall be made available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] T. G. Nguyen, T. V. Phan, D. T. Hoang, H. H. Nguyen, and D. T. Le, "DeepPlace: deep reinforcement learning for adaptive flow rule placement in Software-Defined IoT Networks," *Computer Communications*, vol. 181, pp. 156–163, 2022.

[2] S. Zafar, Z. Lv, N. H. Zaydi, M. Ibrar, and X. Hu, "DSMLB: dynamic switch-migration based load balancing for software-defined IoT network," *Computer Networks*, vol. 214, Article ID 109145, 2022.

[3] W. Rafique, L. Qi, I. Yaqoob, M. Imran, R. U. Rasool, and W. Dou, "Complementing IoT services through software-defined networking and edge computing: a comprehensive survey," *IEEE Communications Surveys & Tutorials*.vol. 22, no. 3, pp. 1761–1804, 2020.

[4] P. Sivakumar, "Improved Resource management and utilization based on a fog-cloud computing system with IoT incorporated with Classifier systems," *Microprocessors and Microsystems*, 2021.

[5] S. Khan, M. Ali, N. Sher, Y. Asim, W. Naeem, and M. Kamran, "Software-defined networks (SDNs) and internet of things (IoTs): a qualitative prediction for 2020," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 11, 2016.

[6] X. Li, L. Ji, H. Zhu, P. Li, X. Jia, and C. Li, "Cellular automata-based simulation of cross-space transmission of energy local area network risks: a case study of a power supply station in Beijing," *Sustainable Energy, Grids and Networks*, vol. 27, Article ID 100521, 2021.

[7] L. Leenen and T. Meyer, "Artificial intelligence and big data analytics in support of cyber defense," *InResearch Anthology on Artificial Intelligence Applications in Security 2021*, IGI Global, , pp. 1738–1753.

[8] T. P. Latchoumi and L. Parthiban, "Quasi oppositional dragonfly algorithm for load balancing in cloud computing environment," 2021.

[9] A. Bicaku, M. Tauber, and J. Delsing, "Security standard compliance and continuous verification for industrial internet of things," *International Journal of Distributed Sensor Networks*, vol. 16, no. 6, Article ID 155014772092273, 2020.

[10] S. Nižetić, P. Šolić, D. López-de-Ipiña González-de-Artaza, and L. Patrono, "Internet of Things (IoT): opportunities, issues and challenges towards a smart and sustainable future," *Journal of Cleaner Production*, vol. 274, Article ID 122877, 2020.

[11] M. Razian, M. Fathian, and R. Buyya, "ARC: anomaly-aware Robust Cloud-integrated IoT service composition based on uncertainty in advertised quality of service values," *Journal of Systems and Software*, vol. 164, Article ID 110557, 2020.

[12] W. T. Vambe, C. Chang, and K. Sibanda, "A review of quality of service in fog computing for the Internet of Things," *International Journal of Flow Control*, vol. 3, no. 1, pp. 22–40, 2020.

[13] J. Ramakrishnan, M. S. Shabbir, N. M. Kassim, P. T. Nguyen, and D. Mavaluru, "A comprehensive and systematic review of the network virtualization techniques in the IoT," *International Journal of Communication Systems*, vol. 33, no. 7, Article ID e4331, 2020.

[14] S. Zroug, I. Remadna, L. Kahloul, S. Benharzallah, and S. L. Terrissa, "Leveraging the power of machine learning for performance evaluation prediction in wireless sensor networks," in *Proceedings of the 2021 International Conference on Information Technology (ICIT)*, pp. 864–869, IEEE, Amman, Jordan, 2021 Jul 14.

[15] D. J. Kennaway, "Trough melatonin levels have no physiological or clinical relevance," *Clinical Psychopharmacology and Neuroscience*, vol. 19, no. 2, pp. 391-392, 2021.

[16] M. Jung and M. Zitterbart, "Cooperative congestion control for cyber-physical systems," in *Proceedings of the 2021 10th Mediterranean Conference on Embedded Computing (MECO)*, pp. 1–4, IEEE, Budva, Montenegro, 2021 Jun 7.

[17] S. Strauss and I. Wonnenberg, "Four short teaching cases on automated decision-making in the public sector," Available at SSRN 3904984, 2021.

[18] M. Venkata Pavan, B. Karnan, and L. Tp, "PLA-Cu reinforced composite filament: preparation and flexural property printed at different machining conditions," *Advanced Composite Materials*, vol. 31, no. 1, pp. 102–117, Article ID 1918608, 2021.

[19] Z. R. Shi, C. Wang, and F. Fang, *Artificial Intelligence for Social Good: A Survey*, arXiv:2001.01818, 2020.

[20] U. Çağlayan, "Performance, energy savings and security: an introduction," in *Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems*, pp. 3–28, Springer, Cham, 2020.

[21] Y. W. Chen, J. P. Sheu, Y. C. Kuo, and N. Van Cuong, "Design and implementation of IoT DDoS attacks detection system based on machine learning," in *Proceedings of the 2020 European Conference on Networks and Communications (EuCNC)*, pp. 122–127, IEEE, Dubrovnik, Croatia, 2020 Jun 15.

[22] W. Jin, R. Xu, T. You, Y. G. Hong, and D. Kim, "Secure edge computing management based on independent microservices providers for gateway-centric IoT networks," *IEEE Access*, vol. 8, pp. 187975–187990, 2020.

[23] S. R. Islam, M. A. Islam, M. A. Kader, and M. M. Kabir, "Design and implementation of a new green IoT gateway module," in *Proceedings of the 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI)*, pp. 1–5, IEEE, Rajshahi, Bangladesh, 2021 Jul 8.

[24] E. Gelenbe, J. Domanska, P. Fröhlich, M. P. Nowak, and S. Nowak, "Self-aware networks that optimize security, QoS, and energy," *Proceedings of the IEEE*, vol. 108, no. 7, pp. 1150–1167, 2020.

[25] W. Tang, B. Li, M. Barni, J. Li, and J. Huang, "An automatic cost learning framework for image steganography using deep reinforcement learning," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 952–967, 2021.

[26] Y. Sun, S. Wang, X. Tang, T. Y. Hsieh, and V. Honavar, "Adversarial attacks on graph neural networks via node injections: a hierarchical reinforcement learning approach," in *Proceedings of the Web Conference 2020*, pp. 673–683, Taipei, China, 2020 Apr 20.

[27] I. Rasheed, F. Hu, and L. Zhang, "Deep reinforcement learning approach for autonomous vehicle systems for maintaining security and safety using LSTM-GAN," *Vehicular Communications*, vol. 26, Article ID 100266, 2020.

[28] A. Rataj, "Random neural networks with hierarchical committees for improved routing in wireless mesh networks with interference," *SN Computer Science*, vol. 1, no. 1, pp. 30–37, 2020.

[29] H. An, Y. Na, H. Lee, and A. Perrig, "Resilience evaluation of multi-path routing against network attacks and failures," *Electronics*, vol. 10, no. 11, p. 1240, 2021.

[30] E. Gelenbe, P. Fröhlich, M. Nowak et al., "IoT network attack detection and mitigation," in *Proceedings of the 2020 9th Mediterranean Conference on Embedded Computing (MECO)*, pp. 1–6, IEEE, Budva, Montenegro, 2020 Jun.

Hindawi

Research Article

AI-Assisted Tuberculosis Detection and Classification from Chest X-Rays Using a Deep Learning Normalization-Free Network Model

Vasundhara Acharya,[1] Gaurav Dhiman [iD],[2,3,4] Krishna Prakasha,[5] Pranshu Bahadur,[5] Ankit Choraria,[5] Sushobhitha M,[6] Sowjanya J,[7] Srikanth Prabhu,[1] Krishnaraj Chadaga,[1] Wattana Viriyasitavat,[8] and Sandeep Kautish [iD][9]

[1]*Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India*
[2]*Department of Electrical and Computer Engineering, Lebanese American University, Byblos, Lebanon*
[3]*Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248002, Manipal, India*
[4]*Department of Project Management, Universidad Internacional Iberoamericana, Campeche, C.P. 24560, Mexico*
[5]*Department of Information Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India*
[6]*Department of Shalakya Tantra, Sri Dharmasthala Manjunatheshwara Institute of Ayurveda and Hospital, Bengaluru, India*
[7]*Department of Kayachikitsa, Sri Dharmasthala Manjunatheshwara Institute of Ayurveda and Hospital, Bengaluru, India*
[8]*Business Information Technlogy Division, Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Bangkok, Thailand*
[9]*Janaki College for Professional Studies, Janakpur, Nepal*

Correspondence should be addressed to Sandeep Kautish; prof.sandeepkautish@gmail.com

Tuberculosis (TB) is an airborne disease caused by *Mycobacterium tuberculosis*. It is imperative to detect cases of TB as early as possible because if left untreated, there is a 70% chance of a patient dying within 10 years. The necessity for supplementary tools has increased in mid to low-income countries due to the rise of automation in healthcare sectors. The already limited resources are being heavily allocated towards controlling other dangerous diseases. Modern digital radiography (DR) machines, used for screening chest X-rays of potential TB victims are very practical. Coupled with computer-aided detection (CAD) with the aid of artificial intelligence, radiologists working in this field can really help potential patients. In this study, progressive resizing is introduced for training models to perform automatic inference of TB using chest X-ray images. ImageNet fine-tuned Normalization-Free Networks (NFNets) are trained for classification and the Score-Cam algorithm is utilized to highlight the regions in the chest X-Rays for detailed inference on the diagnosis. The proposed method is engineered to provide accurate diagnostics for both binary and multiclass classification. The models trained with this method have achieved 96.91% accuracy, 99.38% AUC, 91.81% sensitivity, and 98.42% specificity on a multiclass classification dataset. Moreover, models have also achieved top-1 inference metrics of 96% accuracy and 98% AUC for binary classification. The results obtained demonstrate that the proposed method can be used as a secondary decision tool in a clinical setting for assisting radiologists.

1. Introduction

Lung diseases are often associated with excruciating pain and suffering as they affect the breathing pattern of the patient due to suffocation and related symptoms. Tuberculosis (TB) is one such detrimental variant of lung infections that has created a devastating impact on humankind. Tuberculosis is caused by the *Mycobacterium tuberculosis* bacteria. In general, the lungs are the main target area of this disease, but it can also affect other parts of the body. TB is a contagious disease, i.e., when people infected with TB cough or sneeze, they transmit the disease-causing bacteria in the air. Only a small quantity of these germs are enough to effectively infect a healthy person. Although scientific discoveries and research have been helping to curb the growing influence of TB, the meagre annual medical progress rate in this sector has been unsuccessful in bringing a drastic drop of TB affected patients. According to the Global Tuberculosis Report, 2020, generated by the WHO [1], approximately 10 million people were affected by TB worldwide in 2019. Additionally, HIV/AIDS and TB form a deadly combination. The HIV infection significantly minimizes the strength of the immunity system of an individual, which serves as a favourable condition for an HIV positive patient to contract TB. Out of the 1.4 million deaths caused by TB in 2019, more than 200 thousand patients were HIV positive. In Figure 1, description of chest X-ray of a healthy, viral, and tuberculosis patient is provided [2].

Manual inspection methods are labour intensive and require expertise in that particular domain to give accurate inferences. Therefore, there arises a need to merge the state-of-the-art technological advancements with medical theories and procedures. Artificial intelligence as a sector has boomed magnificently in recent decades and has spread across numerous industries. The entry of artificial intelligence into the medical field has propelled the progress rate tremendously in several types of research and has given scientists the freedom to explore uncharted territories.

Thus, the government authorities and multinational companies in the health sector have been encouraging institutions and academicians to utilize the maximum potential of machine learning and artificial intelligence to accelerate research in important domains such as medical imaging, diagnosis, and drug development. Dermatology, which is an image and screening intensive subfield of medicine, has great synergy with deep learning techniques pertaining to image processing [3–6]. Automation is carried out in electronic health records [7], therapeutic chatbots [8–10], and health monitoring.

Systems [11, 12] have rapidly expanded with the help of deep learning algorithms. With the help of natural language processing, scientists are able to identify and enhance the drug-drug interaction in medical literature [13]. Predictive modelling and decision making are important applications of AI in primary care [14–16].

In TB, computer-aided detection (CAD) is the most widely used artificial intelligence tool. The tool analyses the patient's chest X-rays and determines if the patient is affected by TB. This process reduces the load on the radiologists to meticulously scan through every radiographic film and ultimately speeds up the screening process.

The unique approach in this research is to leverage the power of normalization-free networks to escalate X-ray image classification. We experiment on different normalization-free architectures and compare them with their standard versions to prove their superiority. An AUC of 98% is obtained by the base line model. This proves the potential use of CAD in the diagnosis of the deadly TB.

1.1. Contributions and Related Literature. A number of papers have been published which diagnose TB using deep learning. However, in this article, the novelty lies in a variety of factors from data augmentation and regularization to the use of classification network. Chest Visualization is also discussed in this research which will help the medical personnel. Following are our vital contributions to this research.

(1) The use of RandAugement algorithm for augmentation. It is an automated technique which is known to deliver accurate results.

(2) The use of progressive resizing to augment the images based on their sizes.

(3) Classification using normalization-free network instead of batch normalization to tackle performance issues.

(4) Adaptive grading clipping technique used in this research to avoid the problem of exploding gradients.

(5) Furthermore, the Score-CAM technique used for thorough chest visualization from a medical viewpoint.

A variety of algorithms have been used and compared with the proposed model. There is a drastic improvement in the results obtained by using the above preprocessing and classification techniques. A lot of public datasets have been tested too. Table 1 gives information about the various datasets used in this research. The main novelty in the work is the usage of RandAugment for the augmentation of the data and the utilization of normalization-free network for the classification. Problems with batch normalization: batch normalization is a key component of most image classification models, but it has many undesirable properties. There can be discrepancies between the behavior of the network during training and testing times. While training, the network might have learned and trained to certain batches, which makes the network dependent on that batch-wise setup. So, it might not perform well when a single example is provided at inference. Batch normalization breaks the independence between examples within a batch. This means the examples selected in a batch are significant and lead us to two more prospects: Batch Size matters and inefficient Distributed training that will lead to the network cheating the loss function. To overcome these problems, normalization-free network is utilized and also, the comparison of
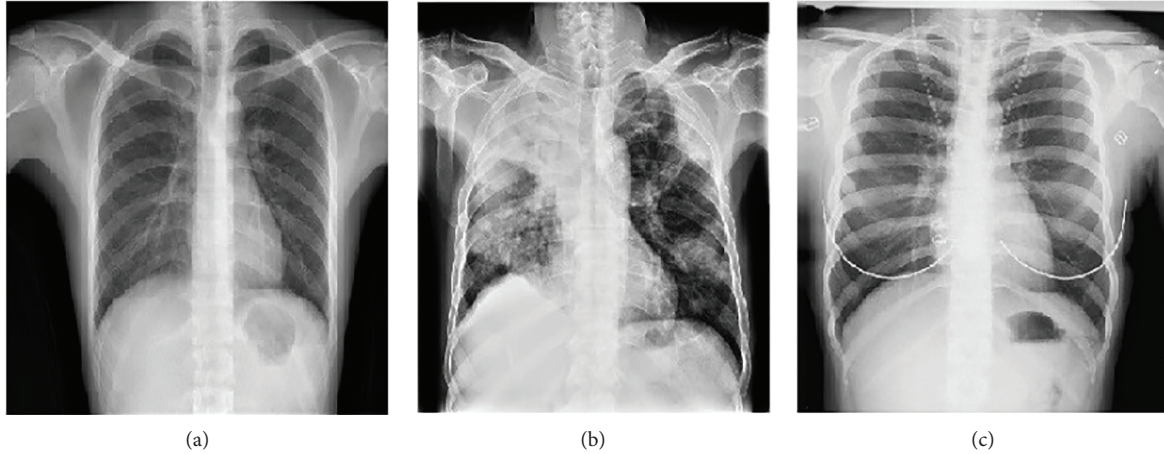
Figure 1: Illustration of Chest X-rays. (a) Healthy. (b) Viral infection. (c) Tuberculosis.

Table 1: Tuberculosis: chest X-ray datasets.

Sr. No	Dataset name/source	Number of images	Reference
1	Find and treat screening program	TB:87, healthy: 113	[17]
2	TB-NEAT research study	TB: 66, Healthy: 134	[17]
3	Montgomery county chest X-ray set	TB:58, Healthy: 80	[18]
4	Shenzhen dataset	TB: 336, healthy: 326	[18]
5	DA	TB: 78, healthy: 78	[19]
6	DB	TB:75, healthy: 75	[19]
7	Belarus tuberculosis portal dataset	TB:304	[20]
8	Tuberculosis (TB) chest X-ray database	TB:3500, healthy: 3500	[21]
9	NIAID TB dataset	TB:3000	[22]
10	JSRT	TB: 154, healthy: 93	[23]
11	KIT	TB: 3828, healthy: 7020	[24]
12	ChestX-ray8	TB: 51,760, healthy: 60,360	[25]
13	TBX11 K	TB: 800, Healthy:3800, Sick: 3800	[26]

NFNet with other state of the art image classification algorithms is mentioned in the paper.

Few recent studies have reported changes in X-ray images in patients at the onset of TB. Here, we review related literature which use AI and deep learning to diagnose this bacterial infection. Liu et al. [2] employed AlexNet and GoogleNet with shuffle sampling to classify TB from chest X-rays and achieved an accuracy of 85.68%. Hooda et al. [27] designed an ensemble of AlexNet, GoogleNet, and ResNet to detect TB. The models were trained from scratch and achieved an accuracy of 88.24%. The benchmark TB dataset (TBX11 K) was proposed by Liu et al. [26]. The authors compared the performance of various deep learning models to achieve simultaneous detection and classification with an accuracy of 88.2%. Hwang et al. [24] designed a modified AlexNet that was pretrained on ImageNet dataset to achieve TB classification. It achieved an accuracy of 90%. The pipeline data augmentation technique and ResNet18 model developed by Ghorakavi et al. [28] achieved an accuracy of 65.77181%. A D-CNN that involved demographic details along with images achieved better accuracy than I-CNN that considered only images in TB detection from chest X-rays [29]. Lakhani and Sundaram et al. [30] employed AlexNet and GoogleNet to detect TB. The ensemble of both gave an AUC score of 0.99. Nguyen et al. [31] proposed that ImageNet weights were insufficient for the modalities like X-Rays and discussed a new technique to obtain low level features by training the models in a multiclass multilabel scenario. Among the models trained, DenseNet-121 outperformed others by achieving an AUC score of 99% and 80% on Shenzhen and Montgomery datasets. Performance of pretrained AlexNet VGG16, VGG-19, and Xception ResNet 50 were compared in [32]. They identified that the features from shallow layers gave better results than deeper layers. Sivaramakrishnan et al. [25] employed InceptionV3 to obtain the classification of TB. They arrived at a conclusion that a supervised deep learning model trained from one population would not have the same diagnostic performance in another population. Hijazi et al. [33] proposed an ensemble of VGG16 and InceptionV3 which utilized the features extracted from original images of chest X-rays and canny edge detected images. The model with probability decision and variation of features led to improved TB detection. Pasa et al. [34] proposed a simple CNN model with five convolutional blocks to achieve the TB classification. They also employed saliency maps and grad-CAMs

visualization techniques and discussed them from the radiologist's perspective. VGG16, Artificial Neural Network (ANN), and a customised CNN were employed to classify between drug resistant and nondrug resistant TB [35]. The ANN outperformed other models as the size of the dataset was small. Vajda et al. [36] employed the atlas-based lung segmentation and feature extraction to obtain the sets of features that could differentiate normal chest.

X-rays from the TB suspicious ones. A neural network-based classifier was utilized to achieve the classification. A maximum area under the curve and accuracy of 99% and 97.03% was obtained with Montgomery and the Shenzhen dataset. An ensemble of classifiers by combining the Support Vector Machines (SVMs) trained using the features extracted from chest X-Rays utilizing GoogLenet, ResNet, and VggNet is proposed in [37]. The models performed extremely well in diagonising TB. In research by Yan et al. [38], ML was used to diagnose TB using CT images.

892 CT scans of patients were included. The overall classification accuracy obtained was 81.09% to 91.05%. The paper concluded that deep learning has a lot of scope to diagnose TB in the future. Deep learning based Mycobacteria detection was conducted in [39]. Two autopsy patients and 40 biopsy cases were used for this research. A 100% specificity was obtained by the algorithms. The sensitivity ranged from 29% to 86%. Podder et al. [40] used transfer learning to diagnose COVID-19 and other diseases. The Modified Xception classifier obtained an accuracy of 84.82%.

The dataset contained chest x-rays of patients. Mondal et al. [41] used an optimized InceptionResNetV2 to diagnose COVID-19. The dataset contained both COVID-19 and non COVID-19 CT images. A maximum accuracy of 96.18% was obtained. A review of various ML and DL algorithms for COVID-19 diagnosis was conducted by Mondal et al. [42]. 52 articles were considered for this extensive review. Results concluded that ResNet-18 and DenseNet 169 were the efficient algorithms. Bharathi et al. [43] developed "CO-ResNet," an optimized algorithm which diagnoses COVID-19 from chest X-rays. A maximum accuracy of 99% was obtained which distinguished COVID-19 from other viral diseases. Bharathi et al. [44] used deep learning to detect lung infections using chest x-rays. Among all the algorithms, VDSNet performed best with 73%. Bharati et al. [45] used CNN and lightGBM to identify the lung carcinoma. ResNet 50 architecture was compared with different models. The metrics used were log loss and ROC curve.

2. Materials and Methods

In this paper, we propose a three-fold method aimed as follows:

(1) Detect if a chest X-ray is related to a healthy patient or to a patient infected with tuberculosis.

(2) To discriminate between sick (but not TB) and TB.

(3) To highlight the affected areas in the chest X-ray symptomatic of TB.

2.1. Dataset Description. The study is carried on two sets. The first set is henceforth referred to as "Set 1" and the second set is henceforth referred to as "Set 2." The model's performance is tested by utilizing the two sets. The first set is referred as "Test Set 1" and the second set is referred as "Test set 2."

Set 1: this set is comprised of the below datasets:

(1) Tuberculosis X-ray (TBX11 K): the dataset [26] contains 11,200 chest X-rays from individual patients of different age groups and genders. Out of the 11,200 images, only 8400 images were provided with ground truths. Hence, they are considered for training and validation. This subset is split into 3800 (Healthy X-rays), 3800 (Sick but not TB X-rays), and 800 (TB X-rays). The images are in PNG format.

(2) Montgomery: this dataset consists of images collected from the *tuberculosis* control program of the Department of Health and Human Services of Montgomery County, MD, USA. 80 X-Rays are normal and 58 X-Rays are abnormal with manifestations of tuberculosis. The images are in DICOM format.

(3) Shenzhen: this dataset consists of X-ray collected by Shenzhen No. 3 Hospital in Shenzhen, Guangdong providence, China. The set consists of 326 normal X-Rays and 336 abnormal X-Rays [18]. The images are in JPEG format.

(4) DatasetA + DatasetB: the images are collected from the National Institute of TB and respiratory diseases, New Delhi, India [19]. Dataset A includes 78 images for each class, showing normal lungs and others with various TB manifestations. Dataset B includes 75 images for each class, showing normal lungs and other TB manifestations.

The images are in DICOM format

(i) Test Set 1:

A separate held out set is utilized. The test set is composed of 1200 Healthy images, 1200 Sick but not TB images, and 2800 images belonging to TB. The set is retained as an online challenge [26]. All the images are in PNG format.

Set 2: this set is comprised of the following datasets:

Belarus, NIAID, and RSNA: the Belarus set [46] includes 306 CXR images belonging to TB positive cases collected from the National Institute of Allergy and Infectious Diseases, Ministry of Health, Republic of Belarus.

The NIAID Set [47] consists of 2800 TB positive images collected from seven different countries. The RSNA set [48] consists of 3,094 normal images collected from the RSNA pneumonia detection challenge.

Test set 2:

Training set 2 is split into a 60-20-20 ratio. This set has 700 images under the normal category and 700 images under the TB category.

2.1.1. Augmentation Techniques and Regularization. To prevent the model from overfitting on the train test split, we utilized 4 layers of RandAugment [49] followed by conversion of the images to Grayscale (with 3 channels, where $r = g = b$) which prevented any colored X-rays after the RandAugments were done. RandAugment is an automated data augmentation method. The main goal of RandAugment is to remove the need for a separate search phase on a proxy task. The search space has two interpretable hyperparameters M and N. N represents the number of augmentation transformations to apply sequentially, and M denotes the magnitude for all the transformations. A number of works enumerated a policy of choosing which transformations to apply out of available transformations, and probabilities for applying each transformation. In this algorithm, to maintain image diversity and reduce the parameter space, learned policies and probabilities for applying each transformation are replaced with a parameter-free procedure of always selecting a transformation with uniform probability 1/K. K represents the number of transformation options. With only these two parameters, RandAugment can be utilized uniformly across different tasks and datasets. It matches or exceeds the performance of the other autoaugment techniques. In the proposed work, a constant value of 0.5 for "M" (magnitude hyper-parameter) is chosen.

Also, the augmentations are applied at every epoch, effectively giving our model new images to train/validate against at each epoch. Augmentations are also applied on the validation set due to the nature of random train/validation split. This resulted in the proposed model to perform better on the test set, than it did on the validation during training. These operations provided very strong explicit regularization. However, to further achieve better model generalization, more techniques are employed. In Figure 2, various augments chosen using RandAugment can be seen.

2.2. Progressive Resizing. The size of an image plays a crucial role in determining the model's performance. Images with smaller dimensions lead to a small network capacity and thus requires lesser regularization. On the other hand, images with larger dimensions require extensive computations and are more prone to be overfitted. Thus, when the model is being trained with variable image sizes, the strength of regularization should be adjusted accordingly to boost the model's accuracy and improve its performance i.e., an image with a smaller dimension works better with weaker regularization and weaker augmentation techniques. Similarly, images with larger dimensions work better with a stronger extent of regularization and augmentation techniques to defeat overfitting. The psuedocode for progressive resizing is described in Algorithm 1.

3. Classification Using Normalization-Free Network

Artificial Intelligence has impacted medical diagnosis in a positive way. The models have been deployed in various hospitals and medical facilities to assist the doctors and radiologists. It also offers a second opinion and stabilizes the conclusion derived from the X-ray images.

The batch normalization technique is used to scale the activations as they pass through the hidden layers and helps in restricting them to a certain range of values. This is achieved by inserting normalization layers after every hidden layer. Despite its several advantages, the batch norm was not the most appropriate alternative as it did not help us in achieving the best possible performance for our specific use case.

One issue always encountered with batch normalization enabled models is the performance discrepancy during training and testing. During the training process, the batch norm technique requires the model to train over numerous batches of preferably large size and the statistics (mean and variance) are computed corresponding to the minibatch.

This tends to make the model batch-dependent. Hence, when fewer images, lesser in comparison to the batch size, are put to test on the model, and the results produced are often inaccurate and deviate from the ground truth. For instance, the batch normalization-enabled model would be successful in training effectively over the X-ray images present in a large batch size of 128, but, when it would be put to test only on a single X-ray image, the statistics (mean and variance) of the test image might vary significantly from the minibatch statistics, which would ultimately lead to erroneous results.

Another issue faced while employing the batch normalization technique is the slower prediction time and extra computation. To tackle the problem of changing ranges of weights between the layers and stabilizing the learning process, the batch norm technique introduces a normalization operation after every hidden layer. Although the desirable results for the above problems are attained, the training time gets compromised. Hence, the normalizing operation after every hidden layer increases the model runtime as well as the computational resources employed by it.

For the reasons stated above, we needed to utilize a network that would not only provide us with robust training performance and highly accurate results but also be time efficient. The normalization-free network achieves significantly better performance than most of its competitors (e.g., EfficientNet-B7) by eliminating the use of batch normalization and slightly modifying the architecture of the normalization-free ResNets [50]. Apart from being lightweight and training efficiently on larger batch sizes, the NFNet utilizes residual branches and Adaptive Gradient Clipping (AGC) that bolster the model's performance [50].

(1) Approach towards residual branches: this component of the NFNet architecture is its "normalization-free" feature. The residual branches of the NFNet architecture make use of 2 scalars, namely α and β [50]. These scalar quantities help in scaling the activations at the beginning and the end of the residual branch, thus restricting the activations to a certain range. This feature of the NFNet is analogous to the scaling operation done by the normalization layers inserted after every hidden layer in the batch

Original Image	Posterize	Rotate
(a)	(b)	(c)
Sharpen	TranslateY	ShearX
(d)	(e)	(f)
ShearY	Solarize	TranslateX
(g)	(h)	(i)
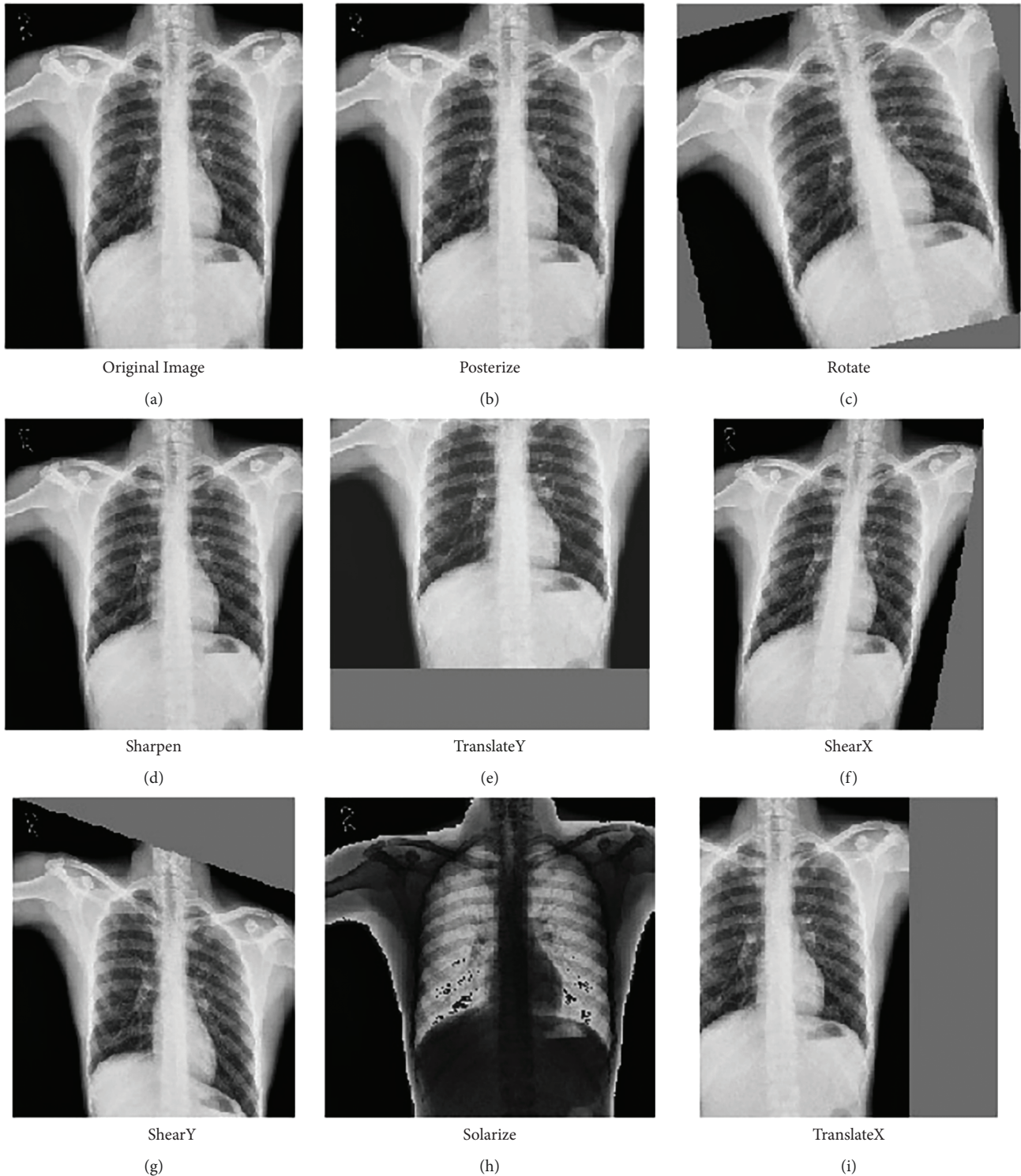
FIGURE 2: Different types of augmentations applied using RandAugment: (a) Original image, (b) Posterize, (c) Rotate, (d) Sharpen, (e) TranslateY, (f) ShearX, (g) ShearY, (h) Solarize, and (i) TranslateX.

normalization enabled models. Hence, the NFNet adopts this merit from the batch normalization technique to compensate for its absence.

(2) Utilization of adaptive gradient clipping: during the backpropagation process, when the norm of the gradient gets bigger as it passes through each layer such that there is an exponential increase in its magnitude, the weights get updated inaccurately and consequently, and the network performance gets hampered. This is known as the exploding gradient issue and it is extremely critical to solving it to improve the overall performance of the model.

The gradient clipping technique is a popular option utilized to tackle exploding gradients. During the training process, when there is a certain norm of a gradient that is very high such that it surpasses a threshold value "lambda" (symbol needs to be inserted), the norm is said to be scaled to the threshold (formula associated with it to be added). Thus, the maximum possible value for a norm of a gradient is the threshold value lambda and any value above the threshold is clipped. For a gradient vector $G = \partial L/\partial \theta$, where L denotes the loss and denotes a vector with all model parameters, the standard clipping algorithm clips the gradient before updating as shown in (1).

$$G = \begin{cases} \dfrac{XG}{\|G\|}, & \text{if } \|G\| > \lambda, \\ \\ G, & \text{otherwise.} \end{cases} \quad (1)$$

However, there are a few drawbacks associated with gradient clipping. The threshold value is very sensitive, and therefore, its appropriate selection is of paramount importance. Secondly, every high gradient jump which surpasses the threshold value is said to be clipped. But, while clipping, the weight associated with the corresponding gradient jump is not taken into account. There might be a few instances where the high gradient jump is justified by a significant corresponding weight, but, since the weight factor is not evaluated, all such gradients are also clipped.

As an improvisation to the existing gradient clipping technique, the adaptive gradient clipping (AGC) takes into account the weight associated with a particular gradient's norm during the filtering process. The AGC is computed as the ratio of the norm of the gradient to the norm of the weight associated with that particular layer. It measures how much the weights are affected for a single gradient descent step. For instance, if the value of the norm associated with a certain gradient is high enough such that it surpasses the threshold, but the weight associated with the same is significant, then the high jump is said to be justified and the gradient is not clipped. However, if the weight associated with the abnormal rise in the norm of the gradient is not sufficient, then that particular gradient is scaled to the threshold value and is said to be clipped. Hence, the weight corresponding to the particular norm of a gradient plays a crucial role in the clipping decision as shown in (2).

$$G_i^l \longrightarrow \begin{cases} G_i^l \longrightarrow \lambda \dfrac{\left\|W_i^l\right\| * F}{\left\|G_i^l\right\|F}, & \text{if } \dfrac{\|G\|F}{\left\|W_i^l\right\| * F} > \lambda, \\ \\ G_i^l, & \text{otherwise,} \end{cases} \quad (2)$$

where $W^l \in \mathrm{R}^{N \times M}$ denotes the weight matrix of the lth layer, $G^l \in \mathrm{R}^{N \times M}$ denotes the gradient with respect to W^l, and the Frobenius norm is denoted by $\|\cdot\|_F$.

Additionally, the AGC serves to be advantageous for our specific use case because it significantly reduces the probability of exploding gradients occurring during the training process by preventing the imprecise updation of the weights. Accurate classification of X-ray images is of prime importance as it is the question of life and death for a particular patient; any slight error can lead to deleterious circumstances. Although batch normalization was not specifically designed to tackle the exploding gradient issue, it was able to mitigate its effects, but not eliminate it. AGC on the other hand was specifically devised to tackle exploding gradient issues and does a much better job in reducing the erroneous effects of gradient explosion as compared to batch norm [51]. The entire flowchart of the proposed model is described in Figure 3. Figure 4 explains the architecture of various deep learning networks.

We denote $h_i \in \mathrm{R}^{m \times n}$ to be the outputs of previous block and $\alpha = 0.2$ be a constant scale factor.

Nontransitional block.

The matrix $X \in {}^{m \times n}$, denotes the output of a nontransitional block.

λ is the composition of operations on h_i that when summed with the original input give the output X i.e.,

$$X = \lambda(hi) + hi. \quad (3)$$

We can denote λ as follows: $\lambda = fk \circ fk-1 \circ \cdots \circ f1(hi)$ where $(f_k: k = 1 \ldots 11)$.

The composition of operations are as follows: (1) $f_1 = (1/\beta)h_i$. Here, input of previous block is scaled by a factor $1/\beta$. Where $\beta_i = {}^P Var(h_i)$ and $Var(h_{i+1}) = Var(h_i) + \alpha^2$. β is the standard deviation of the inputs h_i. (2) For $i = 2 \ldots n - 2$, if i is an even number: fi = ScaledAct (x) Otherwise: fi = WSConv (x).

The scaled activation function [50, 52] (the Gamma Activation) is a ReLU activation scaled by a constant factor $r = \sqrt{2/1 - (1/\pi)}$.

WSConv [52] is the standard weight standardization function: $w_{ij} \longrightarrow w_{ij}'$.

This is done to reparameterize original weights $w_{ij}' \longrightarrow w_{ij} - u_i/\sqrt{N\sigma_I}$,

where $ui = (1/N)Pj \, Wij$ and $\sigma 2 = (1/N)Pj(Wij - ui)2$. (3) f10 = SE(x), which is the queeze and Excecitation block. (4) f11 = αx, (scalar-vector multiplication), where α (as defined earlier) is a scalar value that regulates the rate at which the variance of the activation increases.

Transitional block.

The transitional block is almost identical to the previously described in nontransitional block, with a few changes.

Here, instead of summing outputs of a series of operations $\lambda(h_i)$ with the previous layers input directly, the series of functions is as follows:

$$g(h_i) = WSConv \circ AvgPool \circ ScaledAct\left(\left(\frac{1}{\beta}\right) \times h_i\right), \quad (4)$$

where after scaling the inputs by $(1/\beta)$ feeding the product into the scaled activation (Gamma activation) function.

The matrix is first reduced by average pooling and then expanded again by standard weight standardization.

The output of the transitional block is $Y \in {}^{m \times n} = \lambda(h_i) + G(h_i)$.

The details of the training and validation datasets are described in Table 2 and figuratively given in Figure 5. The
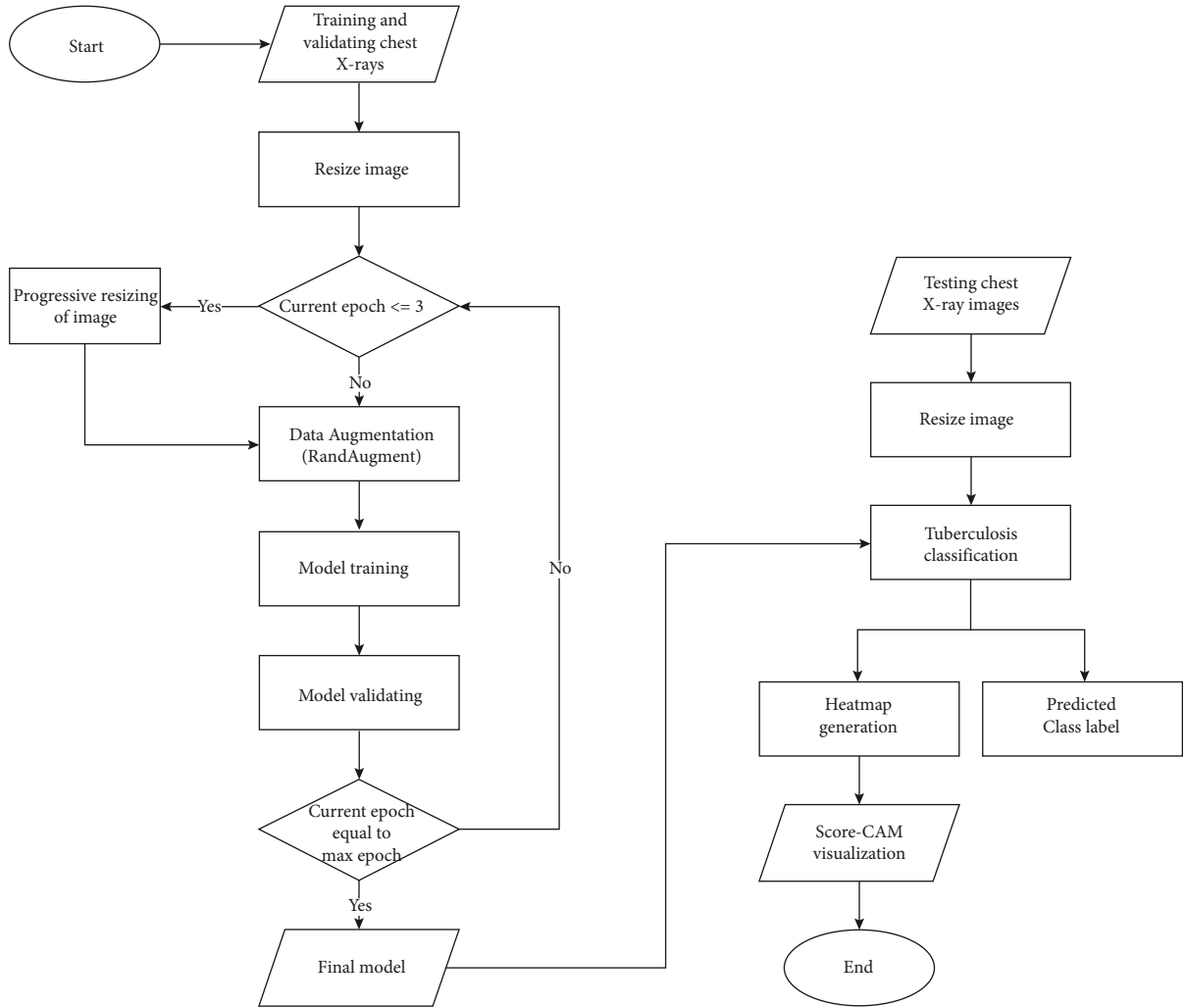
FIGURE 3: Flow of the proposed model.

various training parameters for the NFNet model are explained in Table 3.

4. Performance Criteria for Classification Using Normalization-Free Network

The performance of different models for testing dataset was evaluated after the completion of training and validation phase and was compared using the following performance metrics: Accuracy, Sensitivity, Specificity, Area under 255 Curve (AUC), Average Precision, and Average Recall. TP represents true positive cases, TN represents true negative cases, FN represents false negative cases, and FP represents the number of false positive cases. The metrics are defined below:

(1) Accuracy: the number of positive and negative TB cases identified correctly among all classified cases. It is calculated using the equation given below:

$$\text{Accuracy} = \frac{(TP + TN)}{TP + FN + FP + TN}. \tag{5}$$

(2) Sensitivity: it is the number of positive TB cases identified accurately. It is also called recall. When the number of false negative cases are minimum, the sensitivity is extremely high. It is described using the formula given below:

$$\text{Sensitivity} = \frac{(TP)}{TP + FN}. \tag{6}$$

(3) Specificity: it is the number of negative TB cases identified correctly. When the number of false positives are minimum, the obtained specificity will be maximum. It is calculated using the formula given below:

$$\text{Specificity} = \frac{(TN)}{FP + TN}. \tag{7}$$

(4) AUC (Area Under Curve): the area under the curve when the graph is plotted between the true positive rate and false positive rate. When the AUC is high, it means the model is classifying the instances correctly.
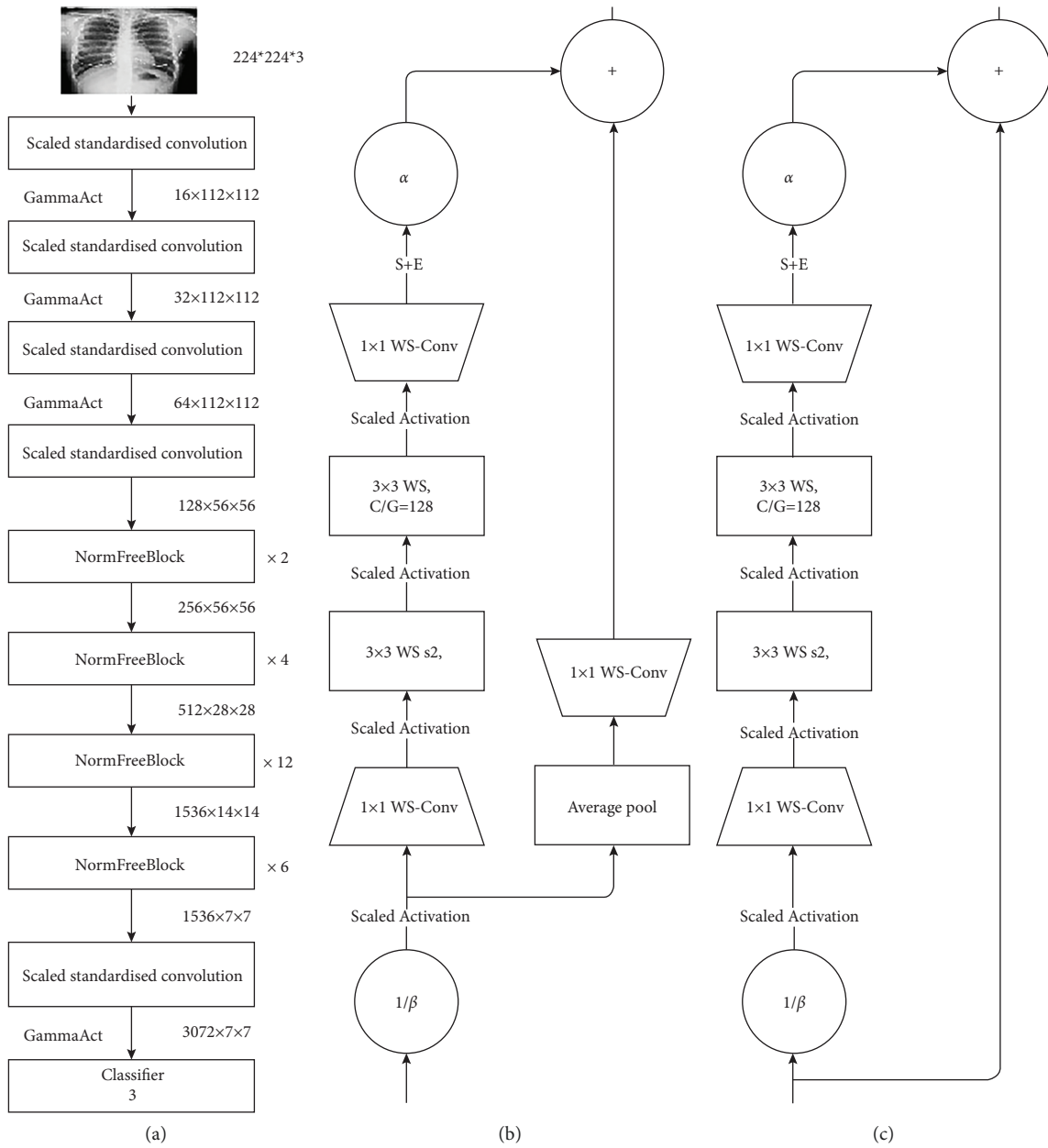
FIGURE 4: (a) Normalization-free network architecture. (b) NFNet transition block. (c) NFNet non transition block.

```
(i)      Require image, resolution, counter, currEpoch
(ii)     if counter-1>0 then
(iii)        image ← resize(image, resolution - 32 X (counter-1))
(iv)         i ← 0
(v)          for i = 4-counter do
(vi)            end
(vii)           image ← randAugment (image)
(viii)       end
(ix)         return image
```

ALGORITHM 1: Progressive resizing.
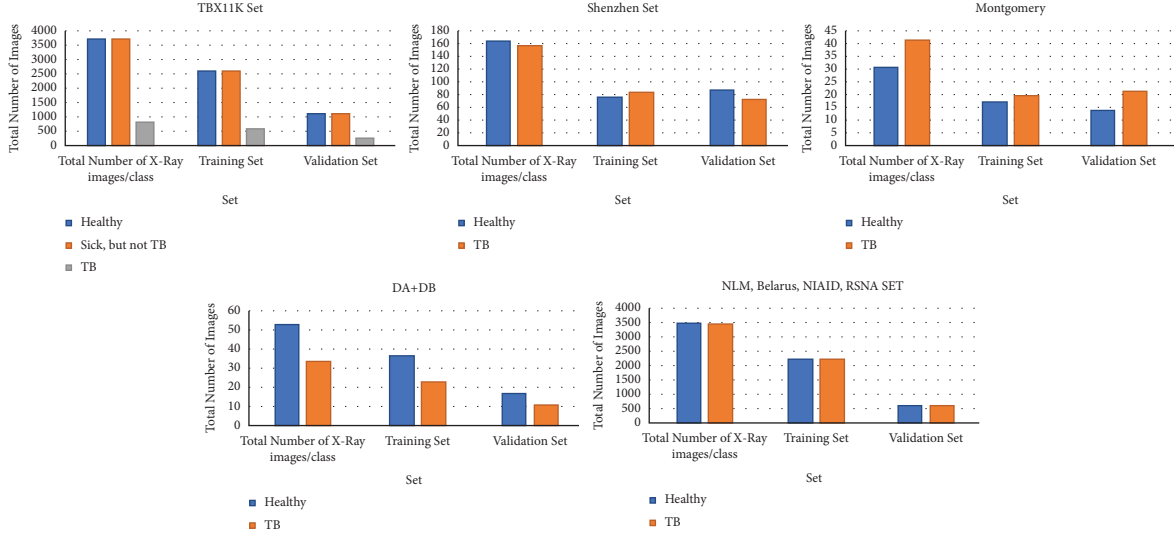
FIGURE 5: Dataset split.

TABLE 2: Details of training and validation set for the classification problem.

Dataset	Type	Total no. of X-Ray images per class	Training set	Validation set
TBX11 K	Healthy	3800	2660	1140
	Sick but not TB	3800	2660	1140
	TB	800	560	240
Shenzhen	Healthy	167	116	51
	TB	160	112	48
Montgomery	Healthy	31	21	10
	TB	42	29	13
DA + DB	Healthy	54	37	17
	TB	34	23	11
NLM (shenzhen and montgomery), Belarus, NIAID, SNA	Healthy	3500	2100	700
	TB	3500	2100	700

(5) Average precision: it summarizes all the values of the precision-recall curve into a value. It also represents the mean of all the precision's obtained. It is calculated using the formula:

$$\text{Average}_{\text{Precision}} = \left(\frac{\sum_n P_n}{n} \right). \qquad (8)$$

(6) Average Recall: the mean of recalls calculated from thresholds of 0.5 to 1 to summarize the distribution. It is calculated using the formula:

$$\text{Average}_{\text{Recall}} = \left(\frac{\sum_n R_n}{n} \right). \qquad (9)$$

4.1. Visualization Technique: Score-CAM. To highlight the areas in the chest X-ray symptomatic of TB, we utilize the Score-CAM [53] technique.

The Score-CAM visualization technique, built on top of the CAM-based visualization method, follows a perturbation 265-based approach. This technique comprises two stages;

the first stage passes the input images into a CNN and generates the activation maps. Subsequently, the maps are upsampled as they are smaller in dimensions as compared to the input image. In the second stage, the activation maps generated are pointwise multiplied with the input image and normalised. The normalization process significantly improves the discriminative ability of the model. The masked inputs are fed to CNN and corresponding scores of the specified target class are generated. This process is repeated 270 until it has been applied to all the generated maps. The Score-CAM metric is given as

$$L_{\text{score-CAM}}^{C} = \text{ReLU}\left(\sum_k \alpha_k^c A_l^k \right), \qquad (10)$$

where

$$\alpha_k^c = C\left(A_l^k \right). \qquad (11)$$

Here, C(.) denotes the channel wise increase in confidence for a particular activation map. The ReLU function is used to eliminate those features which have had no impact
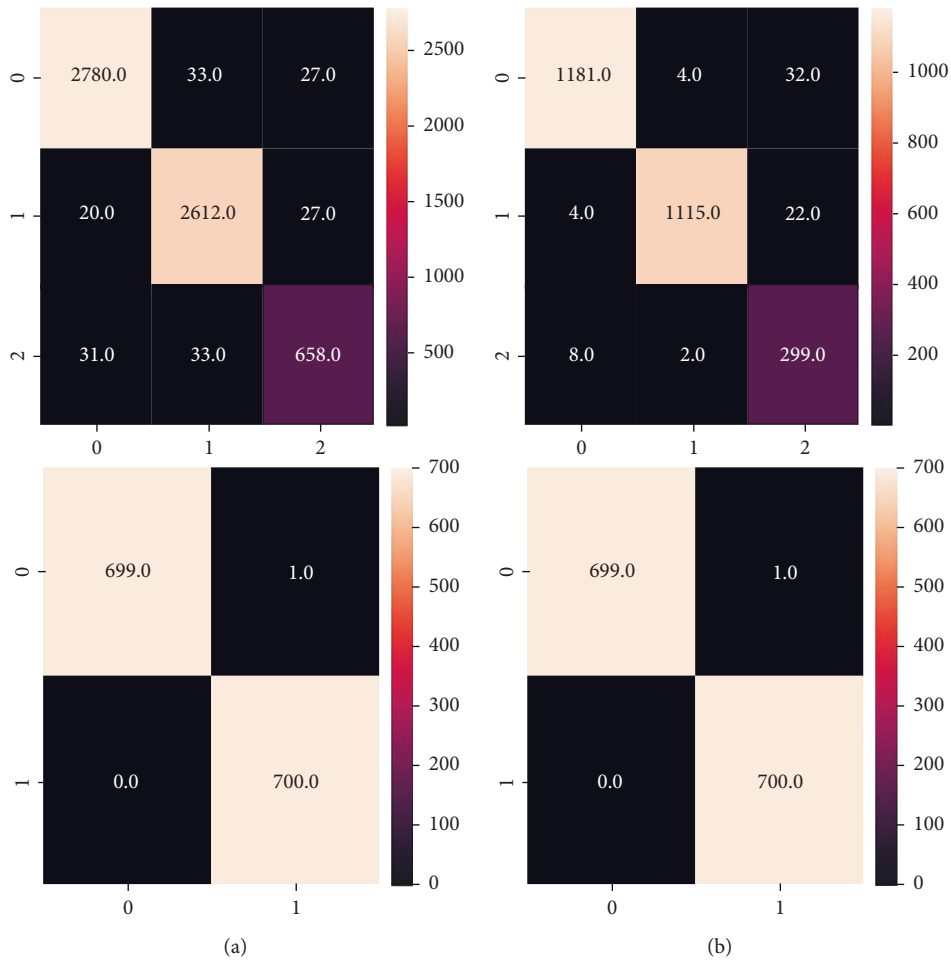
FIGURE 6: (a) Confusion matrices for the training dataset. (b) Confusion matrices for the validation dataset.

on the target class. The proposed work utilizes Score-CAM as it does not require mask sampling or any process for optimisation. Gradients have not been utilized in the course of heat map generation. The removal of global average pooling layer (used in the class activation mapping technique) eliminates the need to retrain the entire process or make any changes to the network structure.

5. Experimental Results and Analysis

In this section, we evaluate the models on the parameters discussed in the previous section. The Score-CAM visualization is also explored in depth. Furthermore, the results are also classified using other deep learning models for TBX11k and Kaggle dataset.

5.1. TB Classification. TB classification using the proposed normalization-free-network model is the main objective of this research. Figure 6 describes the confusion matrix obtained by the two datasets (training + validation). Figure 6(a) represents the confusion matrices for the training data. The matrix on the top represents the TBX11K dataset which consist of three classes namely: healthy, infection (but not

TB), and TB. The bottom matrix is of the Kaggle dataset Test dataset 285 (2) which consists of healthy and TB images. These confusion matrices are obtained after training the normalization-free network model. Figure 6(b) represents the confusion matrices of the above datasets, but on the validation set. As we can see, the false positive and false negative values are extremely low (NonDiagonal elements). This classifier promises us a good accuracy. All the metrics such as accuracy, precision, recall, and others can be effectively calculated using the confusion matrix.

Figure 7 represents the accuracies and losses against the number of epochs for the training and validation sets. From the plots, it can be observed that there is no overfitting since the accuracies and losses are almost similar between the training and test datasets. Further various deep learning architectures were used to classify the datasets along with our proposed model.

The Score-CAM technique was utilized to highlight the regions of the lungs affected by TB. This chest visualization-295 technique can be used to highlight the abnormalities in the specific region in the lungs, aiding the doctor to understand the region of interest. As discussed in the previous section, this process consists of two stages. CNN is used in the beginning to generate activation maps. Furthermore,
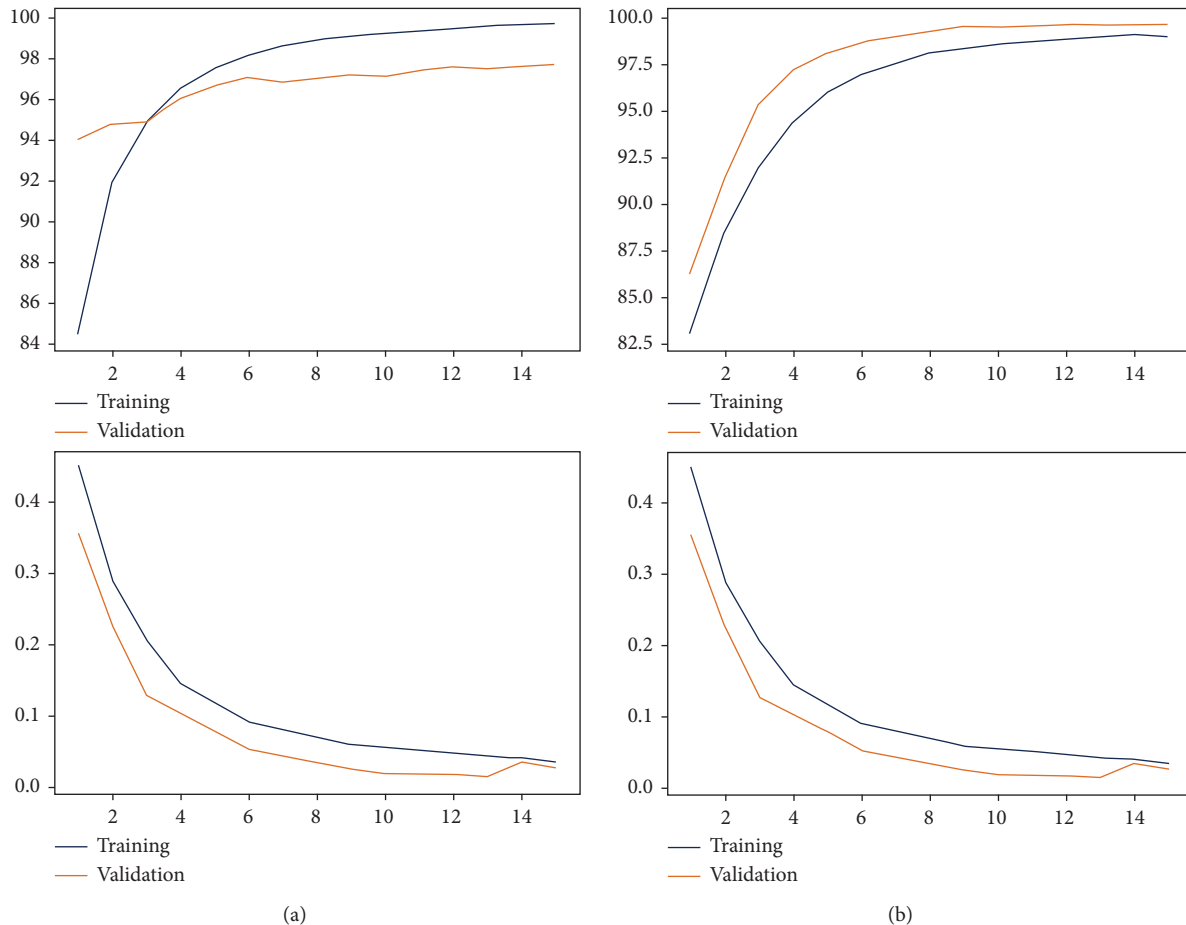
FIGURE 7: (a) Training and validation accuracy and losses versus Epoch for Set 1. (b) Training and validation accuracy and losses versus Epoch for Set 2.

these activation maps are multiplied with the initial image using normalization. The normalization process is extremely important to discriminate between the various image classes. These modified images are again sent to the model for classification. All the generated maps are subjected to 300 of this procedure for accurate analysis. Figure 8 shows the visual analysis of the chest X-ray. The first two rows represent the X-rays of a TB-infected patient. The last two rows are the X-rays of patients infected with other lung diseases (not TB). The first column represents the initial X-ray. The doctors have examined these x-rays and have labelled the region of interest (labelling). This is represented in the second column. The third column represents the marking of the region of interest by the Score-CAM algorithm. From the figure, it is observed that the algorithm identifies the region of interest accurately using a heat map (matches the ground truth given by the domain experts.) However, in some conditions, the heat maps generated can be wrong. This is depicted using Figure 9. The first two rows represent TB cases. The last two rows represent other lung diseases (not TB). The ground truths (labels) by the doctors are present in the second column. From the figure, it is clearly inferred that the generated heat maps are not the same as the ground truths (wrong classification). The Score-CAM is an effective

algorithm, but some false positive and false 310 negative cases are observed.

The data was split into training and testing in the ratio of 80:20. The performance of the models are given in Tables 4 and 5. ResNet-18 is a highly efficient deep learning network which consists of 18 different layers. Millions of images can be easily loaded in this network. It can also classify images into a variety of classes. Furthermore, it is already trained on the ImageNet data. This model obtained an accuracy, AUC, sensitivity, specificity, average 315 precision, and average recall of 91.01%, 95.69%, 76.82%, 96.08%, 89.48%, and 88.75%, respectively, on the TBX11k dataset. For the Kaggle dataset, the obtained accuracy, AUC, sensitivity, specificity, average precision, and average recall are 90%,94%, 72%,95%,87%, and 86.25%, respectively. ResNet50 is a deep residual network which is 50 layers deep. The network is stacked upon each other just like other artificial neural networks. This neural network is pretrained on the ImageNet dataset. The accuracy, AUC, sensitivity, specificity, average precision, and average recall on the TBX11k dataset were 98.76%, 87.02%, 96.99%, 93.67%, and 93.6%, respectively. For the Kaggle dataset, the obtained accuracy, AUC, sensitivity, specificity, average precision, and average recall were 92.61%, 97.6%, 84.02%, 96.88%, 92.41%, and 92.1%, respectively.
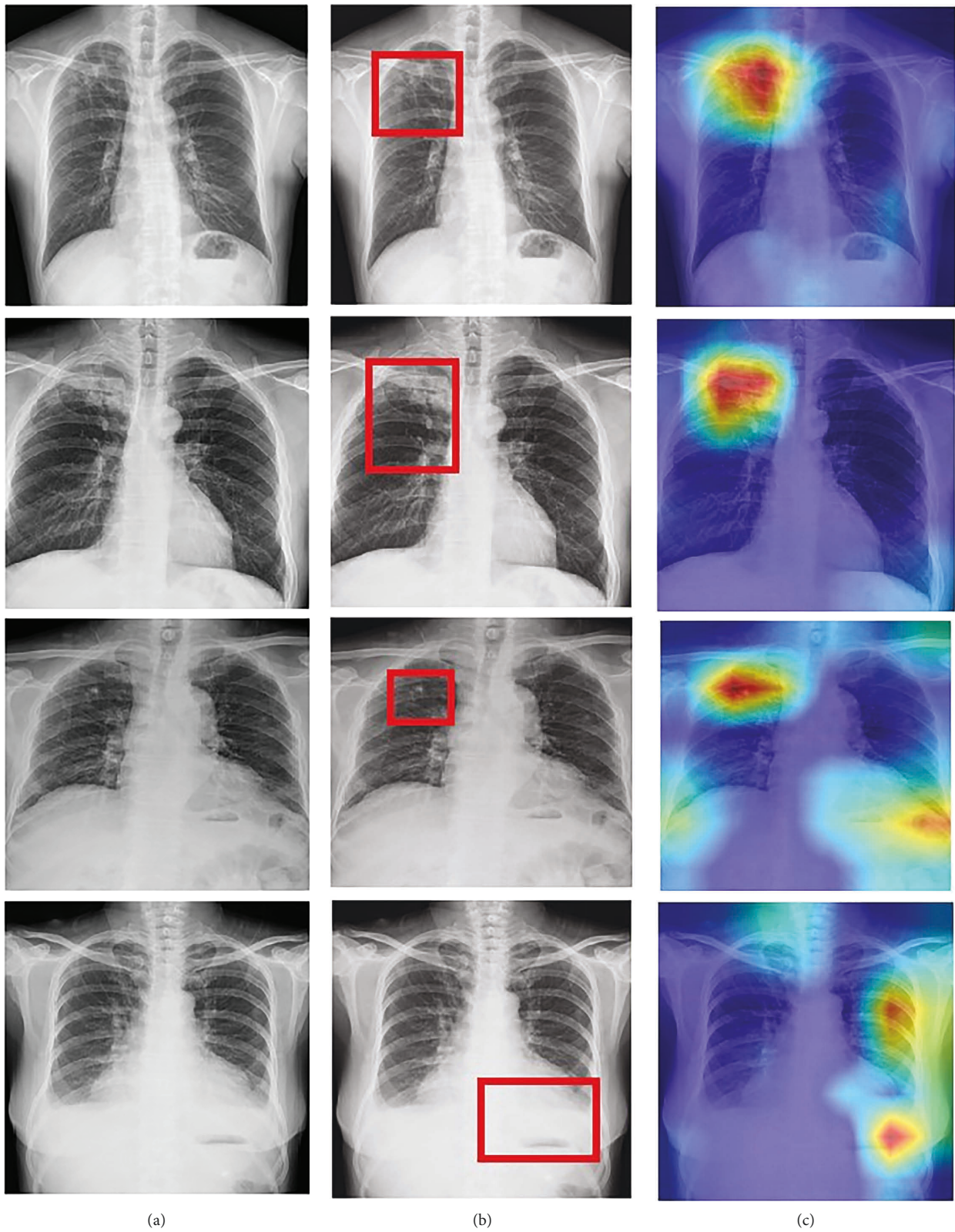
FIGURE 8: Score-CAM visualization of correctly classified TB (first two rows), infected ,and sick but not TB (last two rows) chest X-ray: (a) Original X-ray. (b) Ground truth given by domain expert. (c) Score-CAM heat map generated by the model.
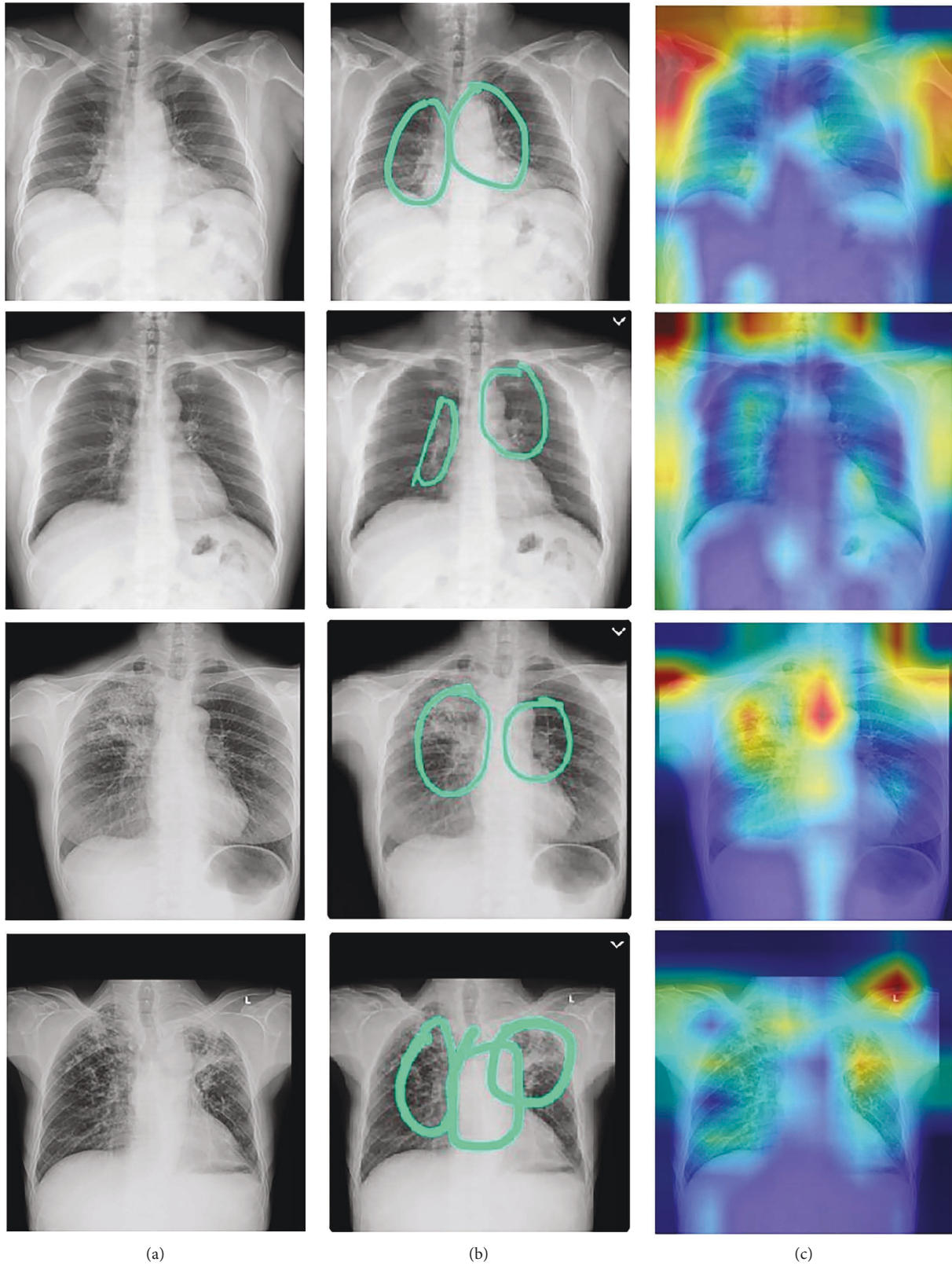
Figure 9: Score-CAM visualization of incorrectly classified actual sick classified as TB (first two rows) and actual TB classified as sick but not TB (last two rows) chest X-ray: (a) Original X-ray. (b) Ground truth given by domain expert. (c) Score-CAM heat map generated by the model.

TABLE 3: Training parameter of model for classification.

Parameter	Value used for training set 1	Value used for training set 2
Batch size	128	32
Learning rate	0.1	0.001
Epochs	15	15
Adaptive gradient clipping	0.16	0.16
Weight decay	1.00E-05	1.00E-05
Optimizer	Stochastic gradient descent	Stochastic gradient descent
StepLR scheduler	Every 2.4 steps	Every 2.4 steps
Gamma	0.97	0.97

TABLE 4: Performance metrics on TBX11K.

Model	Accuracy	AUC (TB)	Sensitivity	Specificity	Average precision	Average recall
ResNet18	91.01	95.69	76.82	96.08	89.48	88.75
ResNet50	94.85	98.76	87.02	96.99	93.67	93.6
DenseNet121	88.04	94.82	71.41	95.89	86.85	85.13
DenseNet201	92.64	97.23	79.91	97.18	91.63	90.6
InceptionV3	89.58	94.95	69.4	97.51	89	86.21
EfficientNet-B7	84.07	73.82	51.78	95.86	82.14	78.39
Proposed model	96.91	99.38	91.81	98.42	96.33	96.1
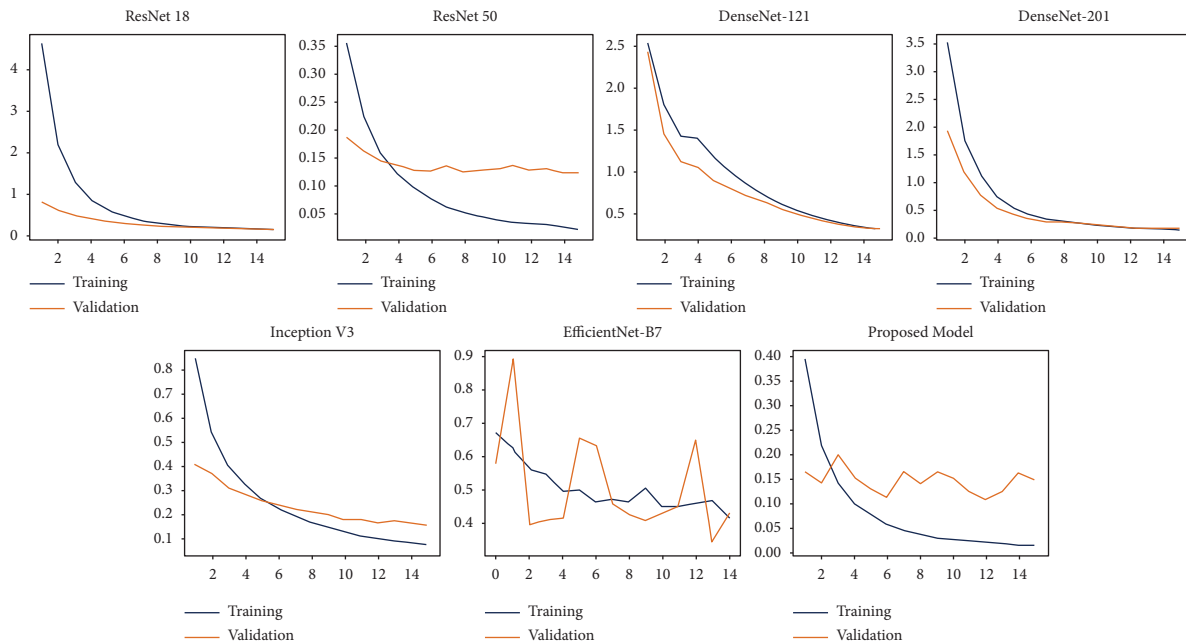
FIGURE 10: Training and validation losses versus Epoch of chest X-ray images present in TBX11k set.

The DenseNet-121 is known to have 120 convolutional layers and four average pooling layers. The weights are spread over multiple inputs for optimal accuracy. These networks were specifically created to counter the vanishing gradient problem. The model obtained an accuracy, AUC, sensitivity, specificity, average precision, and average recall of 98.0%, 94.82%, 71.41%, 95.89%, 86.85%, and 85.13%, respectively. For the Kaggle dataset, the accuracy, AUC, sensitivity, specificity, average precision, and average recall obtained were 88.0%, 94.82%, 71.41%, 95.89%, 86.85%, and 85.13% respectively. Another important CNN model is the DenseNet-201 which consists of 201 layers. It has been already trained by the ImageNet database and can effectively

predict upto 1000 classes. The accuracy, AUC, sensitivity, 330 average precision, and average recall were 92.64%, 97.23%, 79.91%, 97.18%, 91.64%, and 90.6%, respectively, for the TBX11k multiclassification dataset. For the Kaggle dataset, the obtained accuracy, AUC, sensitivity, specificity, average precision, and average recall were 90.11%, 95.14%, 79.81%, 94.21%, 91.22%, and 80.45%, respectively.

InceptionV3 is a deep learning model used to classify images at a high accuracy rate. It includes both symmetric and asymmetric building neurons, including convolutions layer, max pooling, average pooling, drop outs, and fully connected layers. Softmax algorithm is used to compute the model loss. In this network, batch normalization is applied
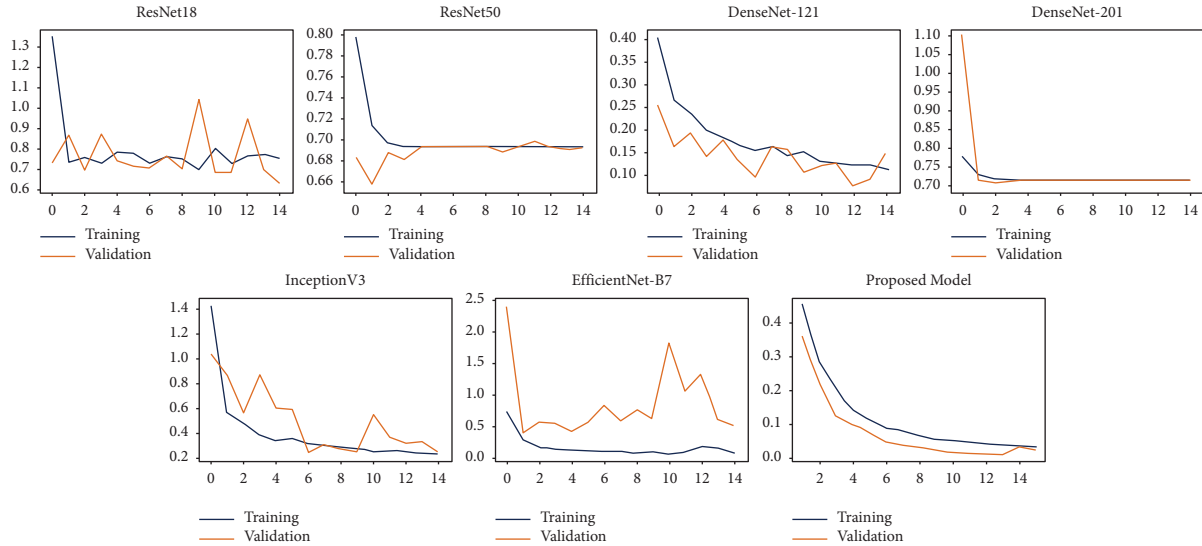
FIGURE 11: Training and validation losses versus Epoch of chest X-ray images present in Kaggle set.

TABLE 5: Performance metrics on Kaggle dataset.

Model	Accuracy	AUC (TB)	Sensitivity	Specificity	Average precision	Average recall
ResNet18	90.23	94.31	72.21	95.41	87.54	86.25
ResNet50	92.61	97.6	84.12	96.88	92.41	92.1
DenseNet121	84.04	93.21	75.43	93.22	84.31	82.11
DenseNet201	90.15	95.26	79.86	94.21	91.22	88.45
InceptionV3	88.62	91.25	80.32	94.31	86.28	84.79
EfficientNet-B7	83.21	74.65	84.32	93.21	80.7	85.22
Proposed model	95.91	98.32	91.78	91.67	95.87	91.90

to the activation inputs. The InceptionV3 was able to obtain an accuracy, AUC, sensitivity, specificity, average precision, and average recall of 89.58%, 94.95%, 69.4%, 97.54%, 89%, and 86.21%, respectively. For the Kaggle dataset, the accuracy, AUC, sensitivity, specificity, average precision, and average recall obtained were 88.62%, 91.25%, 65.21%, 94.31%, 86.28%, and 84.77%, respectively. EfficientNet B7 is a rethinking scaling CNN. It returns an image classification model using transfer learning. This algorithm was developed by a company named AutoML. It also uses a compound coefficient to uniformly scale all the dimensions to its resolution. The obtained accuracy, AUC, sensitivity, specificity, average precision, and average recall were 84.07%, 73.82%, 51.78%, 95.86%, 82.14%, and 78.39%, respectively, for the TBX11k dataset. For the Kaggle dataset, the obtained accuracy, AUC, sensitivity, specificity, average precision, and average recall are 83.21%, 74.65%, 84.32%, 93.21%, 81.31%, and 75.45%, respectively. The proposed normalization-free network was able to perform better than all the models. The accuracy, sensitivity, specificity, average precision, and average recall obtained by the base models were 96.91%, 99.38%, 91.81%, 98.42%, 96.33%, and 96.1%, respectively. This proves that the model is highly efficient in classifying TB. For the binary classification dataset (Kaggle), the obtained accuracy, AUC, sensitivity, specificity, average precision, and average recall are 95.91%, 98.32%,91.78%, 91.67%, 95.33%, and 91.1%. Figure 10 and Figure 11 describe the losses obtained during training and validation for the two datasets.

An outline of comparison with previous literature-based approaches on the TB datasets is compared. The performance is compared with regards to the following measures: accuracy, AUC (TB), sensitivity, specificity, average precision ,and average recall. In research by Li et al. [26], CNN was used to diagnose TB using chest X-rays. AlexNet and GoogleNet were the models used. An accuracy of 85.08% was obtained by the classifiers. Hooda et al. [27] used three architectures: ResNet, GoogleNet, and AlexNet to diagnose TB. The models were further ensembled together to obtain an accuracy and AUC of 88.24% and 93%, respectively. In an article by Liu et al. [26], usage of image-based classifiers for TB diagnosis was considered. The accuracy, sensitivity, specificity, and AUC are 88.2%,88.4%, 89.5% and 93.8%, respectively. Furthermore, TB screening was performed in [23] using deep learning based on chest X-ray images. Three datasets were considered and the AUC obtained by them were 96%, 93%, and 88%, respectively. The TBNet model was created in [28] which diagnosed TB using deep learning models. Mirroring, rotation, and other augmentation techniques were used in the research. However, the ResNet architecture was only able to deliver an accuracy of 81%. Our best model was able to obtain an accuracy, sensitivity, and specificity of 97%, 92%, 99% and respectively. The performance of the comparative researches is described in Table 5.

Tuberculosis is a dangerous disease and it must be identified early to prevent the onset of severe symptoms. AI

TABLE 6: Comparison of various research studies along with our proposed model.

References	Images	Model	Sensitivity	Specificity	AUC score	Accuracy
Liu et al. [2]	4701	Alex net and Google net	—	—	—	85.68
Hooda et al. [27]	1133	Alex net and Google net and ResNet ensemble	88.42	88	93	88.24
Liu et al. [26]	11200	SSD without pretraining and VGG net 1–16 as back bone	88.4	89.5	93.8	88.2
Hwang et al. [24]	10848	Modified pretrained Alex net	—	—	96.4	90
Ghorakavi et al. [28]	800	Reset 18 with data augmentation	—	—	—	65.771
Heo et al. [29]	10000	D-CNN	81.5	96.2	92	—
Lakhani and Sundaram [30]	1007	Ensemble of Alex net and Google Net	97.3	94.7	99	96
Nguyen et al. [31]	800	DenseNet-121	—	—	—	80
Sivaramakrishnan et al. [32]	112,782	Inception v3	72	82	70.54	99
Hijazi et al. [33]	800	VGG 16 and inception V3	90.91	88.64	—	89.77
Pasa et al. [34]	1104	CNN	—	—	92.5	86.2
Jaeger et al. [35]	135	ANN,CNN, VGG16	—	—	66	—
Vajda et al. [36]	814	Neural network	—	—	99	97.03
Lopes et al. [37]	1120	Gooogle net Res Net VGG net SVM	—	—	92.6	84.7
Proposed Model	3500	NF net model	91.81	98.42	99.38	96.91

can be used to diagnose tuberculosis at a faster rate. These models can further help the doctors and radiologists to come to a decision. The classifiers have a tremendous potential in healthcare fields in the near future (Table 6).

6. Conclusion

Tuberculosis is a dangerous bacterial disease which affects the lungs in the human body. It is very important to diagnose this infection early since appropriate treatments can be provided. Computer Aided Diagnosis is a trending topic in Medical Artificial Intelligence. In this research, deep learning-assisted TB diagnosis is performed using normalizer-free network. For augmentation, RandAugment was used to convert the images to gray scale. Further, progressive resizing is used to perform automated preprocessing. Adaptive grading clipping is used to tackle the problem of exploding gradients in this research. A variety of models have been tested and our models achieved an accuracy, AUC, sensitivity, specificity, average precision, and average recall of 98%, 99%, 92%,99%,97%, and 96.1%, respectively. Additionally, a technique called the Score-CAM was used to draw inference from the chest X-rays. Further, the model was compared with the other state-of-art research studies to prove its novelty and superiority. The classifiers can be extremely useful in healthcare and will assist the doctors and medical professionals in performing accurate diagnosis.

In the future, General Adversarial Networks can be utilized to handle data imbalance. This will further boost the efficiency of the model. Chest CT Scan images can be collected from hospitals for computer-aided diagnosis in the near future. The models should be tested on other TB datasets to prove its efficacy. This will make the models more trustworthy. The system can be made user friendly and medical staff can use the models to aid the diagnostic procedure. The deep learning models can also be used to diagnose other diseases such as COVID-19, malaria, liver disease, isochronic heart disease, chronic kidney disease, cancers, and others.

Data Availability

The data used to support the findings of this study are available from author Krishnaraj Chadaga upon request (krishnarajchadaga18@gmail.com).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Sandeep Kautish, Gaurav Dhiman, and Krishna Prakasha proposed the method; Vasund hara Acharya and Krishnaraj Chadaga wrote the paper; Pranshu Agarwal and Ankit Choraria developed the methodology; Wattana Viriyasitavat proof read the paper; Sushobhitha M and Sowjanya annoted the medical images and Srikanth Prabhu analyzed the results.

References

[1] W H O, "Global tuberculosis report 2020," World Health Organization, Geneva, Switzerland, Executive summary, 2020.

[2] C. Liu, Y. Cao, M. Alcantara et al., "Tx-cnn: detecting tuberculosis in chest x-ray images using convolutional neural network," in *Proceedings of the 2017 IEEE international conference on image processing (ICIP)*, pp. 2314–2318, IEEE, Beijing, China, September 2017.

[3] S. Chan, V. Reddy, B. Myers, Q. Thibodeaux, N. Brownstone, and W. Liao, "Machine learning in dermatology: Current applications, opportunities, and limitations," *Dermatologic Therapy*, vol. 10.

[4] S. S. Han, I. J. Moon, W. Lim et al., "Keratinocytic skin cancer detection on the face using region-based convolutional neural network," *JAMA Dermatology*, vol. 156, no. 1, pp. 29–37, 2020, arXiv:.

[5] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[6] M. A. Noyan, M. Durdu, and A. H. Eskiocak, "TzanckNet: a convolutional neural network to identify cells in the cytology of erosive-vesiculobullous diseases," *Scientific Reports*, vol. 10, no. 1, Article ID 18314, 2020.

[7] R. Pivovarov and N. Elhadad, "Automated methods for the summarization of electronic health records," *Journal of the American Medical Informatics Association*, vol. 22, no. 5, pp. 938–947, 2015.

[8] M. Yang, "Painful conversations: therapeutic chatbots and public capacities," *Communication and the Public*, vol. 5, no. 1-2, pp. 35–44, 2020.

[9] K. Rarhi, A. Bhattacharya, A. Mishra, and K. Mandal, "Automated medical chatbot," *SSRN Electronic Journal*.

[10] F. Amato, S. Marrone, V. Moscato, G. Piantadosi, A. Picariello, and C. Sansone, *Chatbots Meet Ehealth: Automatiz- Ing Healthcare*, University of Naples Federico, Italy.

[11] A. Sharma and P. K. Mishra, "Deep learning approaches for automated diagnosis of covid-19 using imbalanced training cxr data," in *Proceedings of the International Conference on Advanced Network Technologies and Intelligent Computing*, pp. 453–472, Springer, Varanasi, India, December 2021.

[12] J. P. Queralta, T. N. Gia, H. Tenhunen, and T. Westerlund, "Edge-ai in lora-based health monitoring," *Fall detection system with fog computing and lstm recurrent neural networks*, vol. 604, 2019.

[13] A. Malawista, X. Wang, M. Trentalange, H. G. Allore, and R. R. Montgomery, "Coordinated expression of tyro3, axl, and mer receptors in macrophage ontogeny," *Macrophage (Houst)*, vol. 3, 2016.

[14] A. Sharma and P. K. Mishra, "Performance analysis of machine learning based optimized feature selection approaches for breast cancer diagnosis," *International Journal of Information Technology*, vol. 14, no. 4, pp. 1949–1960, 2022.

[15] S. Abdul Gafoor, N. Sampathila, S. Ks, and S. K S, "Deep learning model for detection of covid-19 utilizing the chest x-ray images," *Cogent Engineering*, vol. 9, no. 1, Article ID 2079221, 2022.

[16] K. Chadaga, S. Prabhu, K. V. Bhat, S. Umakanth, and N. Sampathila, "Medical diagnosis of covid-19 using blood tests and machine learning," in *Journal of Physics: Conference Series*, vol. 2161, IOP Publishing, Article ID 012017, 2022.

[17] L. Hogeweg, C. I. Sanchez, P. Maduskar et al., "Automatic detection of tuberculosis in chest radiographs using a combination of textural, focal, and shape abnormality analysis," *IEEE Transactions on Medical Imaging*, vol. 34, no. 12, pp. 2429–2442, 2015.

[18] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wang, P.-X. Lu, and G. Thoma, "Two public chest x-ray datasets for´ computer-aided screening of pulmonary diseases," *Quantitative Imaging in Medicine and Surgery*, vol. 4, no. 6, pp. 475–477, 2014.

[19] A. Chauhan, D. Chauhan, and C. Rout, "Role of gist and phog features in computer-aided diagnosis of tuberculosis without segmentation," *PLoS One*, vol. 9, no. 11, Article ID e112980, 2014.

[20] B. P. Health, "Belarus tuberculosis portal," 2021, http://tuberculosis.by/.

[21] T. Rahman, A. Khandakar, M. A. Kadir et al., "Reliable tuberculosis detection using chest x-ray with deep learning, segmentation and visualization," *IEEE Access*, vol. 8, pp. 191586–191601, 2020.

[22] Niaid, "TB portal program dataset," Article ID 45502, 2021, https://data.tbportals.niaid.nih.gov/.

[23] J. Shiraishi, S. Katsuragawa, J. Ikezoe et al., "Development of a digital image database for chest radiographs with and without a lungnodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," *American Journal of Roentgenology*, vol. 174, no. 1, pp. 71–74, 2000.

[24] S. Hwang, H.-E. Kim, J. Jeong, and H.-J. Kim, "A novel approach for tuberculosis screening based on deep convolutional neural networks," in *Medical imaging 2016: Computer-Aided Diagnosis*, vol. 9785, International Society for Optics and Photonics, Article ID 97852W, 2016.

[25] S. Sathitratanacheewin, P. Sunanta, and K. Pongpirul, "Deep learning for automated classification of tuberculosisrelated chest x-ray: dataset distribution shift limits diagnostic performance generalizability," *Heliyon*, vol. 6, no. 8, Article ID e04614, 2020.

[26] Y. Liu, Y.-H. Wu, Y. Ban, H. Wang, and M.-M. Cheng, "Rethinking computer-aided tuberculosis diagnosis," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2646–2655, WA, USA, June 2020.

[27] R. Hooda, A. Mittal, and S. Sofat, "Automated tb classification using ensemble of deep architectures," *Multimedia Tools and Applications*, vol. 78, no. 22, pp. 31515–31532, 2019.

[28] R. S. Ghorakavi, *Tbnet: Pulmonary Tuberculosis Diagnosing System Using Deep Neural Networks*, 2019, https://arxiv.org/abs/1902.08897.

[29] S.-J. Heo, Y. Kim, S. Yun et al., "Deep learning algorithms with demographic information help to detect tuberculosis in chest radiographs in annual workers' health examination data," *International Journal of Environmental Research and Public Health*, vol. 16, no. 2, p. 250, 2019.

[30] P. Lakhani and B. Sundaram, "Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks," *Radiology*, vol. 284, no. 2, pp. 574–582, 2017.

[31] Q. H. Nguyen, B. P. Nguyen, S. D. Dao et al., "Deep learning models for tuberculosis detection from chest x-ray images," in *Proceedings of the 2019 26th International Conference on Telecommunications (ICT)*, pp. 381–385, IEEE, Hanoi, Vietnam, April 2019.

[32] R. Sivaramakrishnan, S. Antani, S. Candemir et al., "Comparing deep learning models for population screening using chest radiography," in *Medical Imaging 2018: ComputerAided Diagnosis*, vol. 10575, International Society for Optics and Photonics, Article ID 105751E, 2018.

[33] M. H. A. Hijazi, S. Kieu Tao Hwa, A. Bade, R. Yaakob, and M. Saffree Jeffree, "Ensemble deep learning for tuberculosis detection using chest x-ray and canny edge detected images," *IAES International Journal of Artificial Intelligence*, vol. 8, no. 4, p. 429, 2019.

[34] F. Pasa, V. Golkov, F. Pfeiffer, D. Cremers, and D. Pfeiffer, "Efficient deep network architectures for fast chest x-ray tuberculosis screening and visualization," *Scientific Reports*, vol. 9, no. 1, pp. 6268-6269, 2019.

[35] S. Jaeger, O. H. Juarez-Espinosa, S. Candemir et al., "Detecting drug-resistant tuberculosis in chest radiographs," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 12, pp. 1915–1925, 2018.

[36] S. Vajda, A. Karargyris, S. Jaeger et al., "Feature selection for automatic tuberculosis screening in frontal chest radiographs," *Journal of Medical Systems*, vol. 42, no. 8, pp. 146–211, 2018.

[37] U. Lopes and J. F. Valiati, "Pre-trained convolutional neural networks as feature extractors for tuberculosis detection," *Computers in Biology and Medicine*, vol. 89, pp. 135–143, 2017.

[38] C. Yan, L. Wang, J. Lin et al., "A fully automatic artificial intelligence–based ct image analysis system for accurate detection, diagnosis, and quantitative severity evaluation of pulmonary tuberculosis," *European Radiology*, vol. 32, no. 4, pp. 2188–2199, 2022.

[39] Y. Zaizen, Y. Kanahori, S. Ishijima et al., "Deep-learning-aided detection of mycobacteria in pathology specimens increases the sensitivity in early diagnosis of pulmonary tuberculosis compared with bacteriology tests," *Diagnostics*, vol. 12, no. 3, p. 709, 2022.

[40] P. Podder, S. Bharati, M. Mondal, and A. Khamparia, "Rethinking the transfer learning architecture for respiratory diseases and covid-19 diagnosis," in *Biomedical Data Analysis and Processing Using Explainable (XAI) and Responsive Artificial Intelligence (RAI)*, pp. 105–121, Springer, Berlin, Germany, 2022.

[41] M. R. H. Mondal, S. Bharati, and P. Podder, "Co-irv2: optimized inceptionresnetv2 for covid-19 detection from chest ct images," *PLoS One*, vol. 16, no. 10, Article ID e0259179, 2021.

[42] M. R. H. Mondal, S. Bharati, and P. Podder, "Diagnosis of covid-19 using machine learning and deep learning: a review," *Current Medical Imaging Formerly Current Medical Imaging Reviews*, vol. 17, no. 12, pp. 1403–1418, 2021.

[43] S. Bharati, P. Podder, M. R. H. Mondal, and V. S. Prasath, "Co-resnet: optimized resnet model for covid-19 diagnosis from x-ray images," *International Journal of Hybrid Intelligent Systems*, vol. 17, no. 1-2, pp. 71–85, 2021.

[44] S. Bharati, P. Podder, and M. R. H. Mondal, "Hybrid deep learning for detecting lung diseases from x-ray images," *Informatics in Medicine Unlocked*, vol. 20, Article ID 100391, 2020.

[45] S. Bharati and P. Podder, "Performance of cnn for predicting cancerous lung nodules using lightgbm," *Artificial Intelligence for Data-Driven Medical Diagnosis*, vol. 1, pp. 1–18.

[46] B. P. Health, 2020, http://tuberculosis.by/.

[47] A. Rosenthal, A. Gabrielian, E. Engle et al., "The tb portals: an open-access, web-based platform for global drug-resistant-tuberculosis data sharing and analysis," *Journal of Clinical Microbiology*, vol. 55, no. 11, pp. 3267–3282, 2017.

[48] Kaggle, "RSNA pneumonia detection challenge," 2018, https://www.kaggle.com/c/rsna-pneumonia/.

[49] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, "Randaugment: practical automated data augmentation with a re- duced search space," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 702-703, WA, USA, June 2020.

[50] A. Brock, S. De, S. L. Smith, and K. Simonyan, "High-performance large-scale image recognition without normalization," 2021, https://arxiv.org/abs/2102.06171.

[51] T. Chen, Z. Zhang, X. Ouyang, Z. Liu, Z. Shen, and Z. Wang, "Training binary neural networks without batch normalization," 2021, https://arxiv.org/abs/2104.08215.

[52] A. Brock, S. De, and S. L. Smith, "Characterizing signal propagation to close the performance gap in unnormalized resnets, CoRR abs/2101," Article ID 08692, https://arxiv.org/abs/2101.08692.

[53] H. Wang, Z. Wang, M. Du et al., "Score-cam: score-weighted visual explanations for convolutional neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 24-25, WA, USA, June 2020.

Research Article

The Drosha-Independent MicroRNA6778-5p/GSK3β Axis Mediates the Proliferation of Gastric Cancer Cells

Mingjun Ren ,[1,2] **Li Xing,**[3] **Wanping Wang,**[3] **Wanying Bi,**[3] **Wanjun Wu,**[4] **Gui Jiang,**[4] **Weiji Wang,**[5] **Xingdong Liang,**[1,2] **Manran Liu,**[6] **and Shifu Tang** [1,2]

[1]*Department of Laboratory Medicine, Liuzhou People's Hospital, Liu Zhou 545006, China*
[2]*Liuzhou Key Laboratory of Precision Medicine for Viral Diseases, Liu Zhou 545006, China*
[3]*Graduate School, Guangxi University of Chinese Medicine, Nanning 530000, China*
[4]*Department of Laboratory Medicine, Liuzhou Traditional Chinese Medicine Hospital, Liu Zhou 545006, China*
[5]*Gastrointestinal Surgery, Liuzhou People's Hospital, Liu Zhou 545006, China*
[6]*Laboratory Medical College, Chongqing Medical University, Chongqing 400016, China*

Correspondence should be addressed to Shifu Tang; sftang2019@163.com

Background. Gastric cancer (GC) is a primary cause of cancer death around the world. Previous studies have found that Drosha plays a significant role in the development of tumor cells. Soon after, we unexpectedly found that the expression of micro-RNA6778-5p (miR6778-5p) is unconventionally high in the gastric cancer cells low-expressing Drosha. So, we designed the Drosha interference sequence and recombined it into a lentiviral vector to construct Drosha knockdown lentivirus and transfected the Drosha knockdown lentivirus into gastric cancer cells to establish Drosha knockdown gastric cancer cell lines. We aimed to explore the effect of microRNA6778-5p on the proliferation of gastric cancer cells with Drosha knockdown and its intrinsic mechanism. *Methods.* We designed the Drosha interference sequence and recombined it into a lentiviral vector to construct Drosha knockdown lentivirus and transfected the Drosha knockdown lentivirus into gastric cancer cells to establish Drosha knockdown gastric cancer cell lines. After transfecting miR6778-5p mimics and inhibitor into gastric cancer cell lines with Drosha knockdown, the expression levels of miR6778-5p mimics in Drosha low-expressing gastric cancer cells increased, while miR6778-5p inhibitor decreased the expression levels of miR6778-5p. The Cell Counting Kit-8 (CCK-8) experiment was used to detect the proliferation ability of gastric cancer cells after overexpression or knockdown of miR6778-5p and bioinformatics predicted the relationship between miR6778-5p and glycogen synthase kinase-3β (GSK3β). *Results.* After infection with the Drosha knockdown lentivirus, Drosha's mRNA and protein levels were significantly downregulated in gastric cancer cells. The expression levels of miR6778-5p mimics in Drosha low-expressing gastric cancer cells increased, while miR6778-5p inhibitor decreased the expression levels of miR6778-5p. Overexpression of miR6778-5p significantly enhanced the proliferation ability of Drosha low-expression gastric cancer cells; on the contrary, knocking down miR6778-5p weakened the proliferation ability of Drosha low-expression gastric cancer cells. Bioinformatics predicted that miR6778-5p targeted glycogen synthase kinase-3β (GSK3β) and the mRNA and protein levels of GSK3β decreased significantly after overexpression of miR6778-5p. *Conclusion.* miR6778-5p promotes the proliferation of Drosha low-expressing gastric cancer cells by targeting GSK3β.

1. Introduction

Gastric cancer is a malignant tumor with high morbidity and mortality worldwide [1]. There are approximately 110000 new cases and 770000 deaths annually, accounting for 5.6% in total number of cancers and 7.7% in total number of cancer deaths [2]. Although some progress has been made in

clinical therapy, the long-term survival for patients with advanced cancer remains poor, with a five-year survival rate of only 12% [3]. Among them, the recurrence of GC is a principal element affecting the survival rate of advanced patients. So, it is urgent to explore the underlying molecular mechanism of the development process of GC and is beneficial to discover new therapy targets of gastric cancer.

Drosha belongs to the endoribonuclease III superfamily. As a core nuclease, Drosha performs the initial step of miRNA processing by cutting the hairpin structure embedded in the primary transcript in the nucleus [4]. Drosha is significant for microRNA processing, which can convert the initial microRNA (pri-miRNA) into the precursor microRNA (pre-miRNA) in the nucleus, laying foundation for the maturation of microRNA [5]. Drosha's RNA interference leads to massive accumulation of pri-miRNA in vivo, with consequent reduction in pre-miRNA and mature miRNA. In addition, there is also a mirtron pathway that does not rely on Drosha to cause miRNA biogenesis, which can suppress or promote cancer [6, 7]. Indeed, previous research has found that Drosha plays an extremely vital part in tumor cell development, and the abnormal expression of Drosha will lead to changes in the expression of microRNA [8, 9]. Especially in some Drosha knockout models, some miRNA expressions are upregulated instead [10]. As we all know, microRNAs are a class of highly conserved and short-stranded non-coding RNAs that partake in the regulation of cell multiplication, apoptosis, and metabolism, and their abnormal expression also affects the progression of tumors [11–13]. The currently recognized mechanism of microRNA's influence on tumors is through targeted inhibition of its related genes to perform their functions [14, 15]. Our preliminary research found that Drosha knockdown in GC cells reduces the migration of cancer cells, accompanied by upregulation or downregulation of miRNA, which is a poor prognostic factor. Of note, we unexpectedly found that microRNA6778-5p (miR6778-5p) expression is unconventionally high in the gastric cancer cells low-expressing Drosha, while miR6778-5p is a non-canonical miRNA type that does not depend on Drosha, spliced from the SHMT1 intron [16]. To clarify the molecular mechanism of miR6778-5p regulating the low expression of Drosha in gastric cancer, we used the miRTarBase and TargetScan databases to find that GSK3β was a potential target gene of miR6778-5p. c-Myc, a class of target genes, is upregulated to accelerate cell multiplication by β-catenin. Meanwhile, GSK3β takes a pivotal effect on controlling cell proliferation by negatively regulating the transcription activity of β-catenin [17]. To our knowledge, till now miR6778-5p/GSK3β axis has no relevant reports on cell multiplication. It is unclear that whether miR6778-5p has an essential part via GSK3β in accelerating GC cell multiplication with Drosha knockdown.

In this study, we found that mir-6778-5p played a positive role in GC cell proliferation. The upregulation of miR6778-5p results from Drosha interference in GC cells. The mimics and inhibitors of miR6778-5p can enhance or inhibit the multiplication of GC cells low-expressing Drosha. Moreover, miR6778-5p regulates the multiplication behavior of GC cells low-expressing Drosha by targeting GSK3β. In short, the microRNA6778-5p/GSK3β axis mediates the multiplication of GC cell lines low-expressing Drosha.

2. Methods

2.1. Cell Lines and Culture Conditions.
The GC cell lines MGC-803 and SGC-7901 needed for the experiment were presented by Professor Yang Ke from Beijing Cancer Institute (Beijing, China). These cells were incubated in RPMI1640 (Gibco, USA) or DMEM (Gibco, USA) supplemented with 10% FBS (Gibco, USA). Cells were cultured in humidified air at 37°C containing 5% CO2.

2.2. Transfection Assay.
The cells were placed in 2 ml whole culture medium with 1×10^5 cells per well in in 6-well plates, respectively. Infected cells with the lentivirus contain Drosha-shRNA sequence (5′-AACGAGUAGGCUUCGU-GACUU-3′) or negative control sequence (5′-UUCUCC-GAACGUGUCACGUTT-3′) (GenePharma Co., Ltd. Shanghai) when the cell density reached 60%. Furthermore, they were screened with 1 μg/ml puromycin, establishing MGC-803/Drosha KD or SGC-7901/Drosha KD cell lines, respectively. Similarly, when reaching 60% confluence in 6-well plates, cells were treated with 100 pmol miR6778-5p mimics, 100 pmol miR6778-5p inhibitor, or the corresponding random sequence RNA Oligo (negative control) (GenePharma Co., Ltd. Shanghai) separately according to the protocol of Lipofectamine 2000 transfection kit (Invitrogen, USA).

2.3. RNA Extraction and qRT-PCR Detection.
Total RNA was isolated from the corresponding gastric cancer cells using Trizol (Invitrogen, USA). RNA quantity was detected by spectrophotometry and by agarose gel electrophoresis. qRT-PCR was done using the PrimeScript RT reagent kit and SYBR Premix Ex TaqTM (Takara, Japan) following the manufacturer's instruction. Two-step amplification, Holding stage: 95°C, 30 s (1 cycle); Cycling stage: 95°C, 5 s–60°C, 34 s (40 cycles); Melt curve stage: 95°C, 15 s–60°C, 60 s–95°C, 15 s (1 cycle). β-Actin was used as the internal control and primer sequences were: Drosha: 5′-CGATGATGCAGG-GAAACACATG-3′ (forward) and 5′-TTATTTCTTGATGT CTTCAGTCT-3′ (reverse). miR6778-5p: 5′-GCGAGTGG-GAGGACAGGAG-3′ (forward) and 5′-ATCCAGTGCAG GGTCCGAGG-3′ (reverse). GSK3β: 5′-CCTTAACCTGC TGCTGGACT-3′ (forward) and 5′-AGCTCTGGTGCCCT GCCAGAT-3′ (reverse). β-Actin: 5′-AGGGGCCGGACTC GTCATACT-3′ (forward) and 5'-GGCGGCACCACCATG TACCCT-3′ (reverse).

2.4. Western Blot.
Western blot assays were done as described previously [18]. Cells were harvested with RIPA buffer (Beyotime, China) to extract total protein measured with the BCA protein assay kit (Beyotime, China), isolated by 10% SDS-PAGE gel electrophoresis. The protein is transferred to the PVDF membrane. After incubating with primary antibodies overnight at 4°C, the corresponding HRP-conjugated secondary antibodies (Beyotime, China)
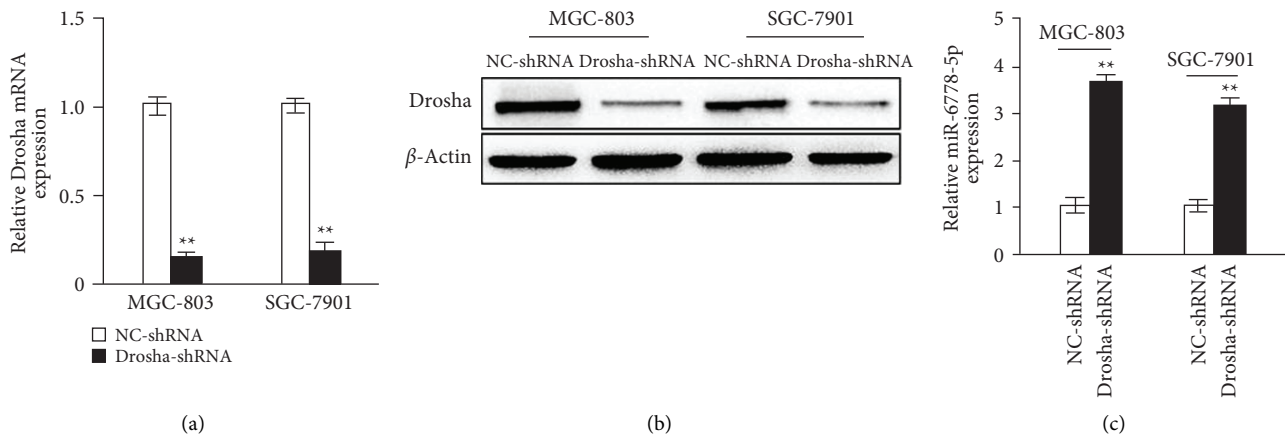
FIGURE 1: Drosha-independent miR6778-5p expression increases in gastric cancer cells. (a) The mRNA expression levels of Drosha were determined by qRT-PCR in gastric cell lines with Drosha knockdown ($^{**}P < 0.05$). (b) The protein expression levels of Drosha were verified by WB in gastric cell lines with Drosha knockdown ($^{**}P < 0.05$). (c) The expression levels of miR6778-5p were detected by qRT-PCR in gastric cancer cells low-expressing Drosha ($^{**}P < 0.05$).

were subsequently applied and immunodetection was performed using the enhanced chemiluminescence system (Cool-Imager).

2.5. Cell Multiplication Assay.
The cell multiplication rate was detected by the CCK-8 assay (Beyotime, China). Cells were placed on 96-well plates with 3×10^3 cells per well. Cells were processed by 100 pmol miR6778-5p mimics and 100 pmol miR6778-5p inhibitor (GenePharma Co., Ltd., Shanghai) at different points in time when 60% confluence was reached. After the treatment, the medium was discarded. A small orifice was placed at 10 μL CCK-8 reagent and continued to incubate for 4 hours before adding DMSO (200 μl per well), and shaken slowly on a shaker for 10 min. Measurement was performed by using an absorbance meter.

2.6. Statistical Analysis.
Statistical analysis was performed using the SPSS standard version 19.0 software. The data were presented as mean ± standard deviation ($\overline{x} \pm s$). Each experiment was repeated at least 3 times. Pairwise comparisons between groups were performed using the LSD t-test, and $P < 0.05$ was considered statistically significant.

3. Results

3.1. The Expression of miR6778-5p Increases in GC Cells with Drosha Knockdown.
To investigate the role of miR6778-5p in GC low-expressing Drosha, we constructed a vector lentivirus with the Drosha interference sequence and then infected MGC-803 GC cell and SGC-7901 GC cell with the virus to establish GC cell lines with Drosha low expression. The mRNA and protein expression of Drosha were significantly reduced compared with the control group (Figures 1(a) and 1(b)) and the miR6778-5p expression was visibly increased (Figure 1(c)) in MGC-803 and SGC-7901 cells transfected with the Drosha knockdown lentivirus.

3.2. Overexpression of miR6778-5p Promotes the Multiplication of GC Cells Low-Expressing Drosha.
To demonstrate whether miR6778-5p regulates the multiplication of GC cells low-expressing Drosha, we transfected Drosha low-expressing GC cells (MGC-803/Drosha KD and SGC-7901/Drosha KD) with miR6778-5p mimics and evaluated the proliferation potential change of MGC-803/Drosha KD and SGC-7901/Drosha KD treated with miR6778-5p mimics. As shown in Figures 2(a) and 2(b), the expression levels of miR6778-5p increased significantly and enhanced the proliferation ability of MGC-803/Drosha KD and SGC-7901/Drosha KD cells treated with miR6778-5p mimics. These results showed that miR6778-5p has a positive influence on multiplication of GC cells low-expressing Drosha.

3.3. Knockdown of miR6778-5p Inhibits the Multiplication of Low-Expression Drosha GC Cells.
In order to further verify whether miR6778-5p controls the proliferation ability of Drosha low-expressing GC cells, we transferred the inhibitor of miR6778-5p to lower expression Drosha GC cells and then evaluated the proliferation potential change of MGC-803/Drosha KD and SGC-7901/Drosha KD treated with miR6778-5p inhibitor. Expectedly, the expression levels of miR6778-5p decreased and the proliferation ability reduced significantly in MGC-803/Drosha KD and SGC-7901/Drosha KD cell lines after treating with miR6778-5p inhibitor (Figures 3(a) and 3(b)). It is confirmed once again that miR6778-5p promotes the multiplication capacity of GC cells low-expressing Drosha.

3.4. The miR6778-5p/GSK3β Axis Mediates the Multiplication Potential of Drosha Low-Expressing GC Cells.
As is well-known, microRNA regulates human tumor development by binding to the 3'UTR end of its target genes to inhibit the expression of coding target genes. We searched the microRNA target gene databases to find that GSK3β was a potential target gene of miR6778-5p (Figure 4(a)). In order

(a)   (b)
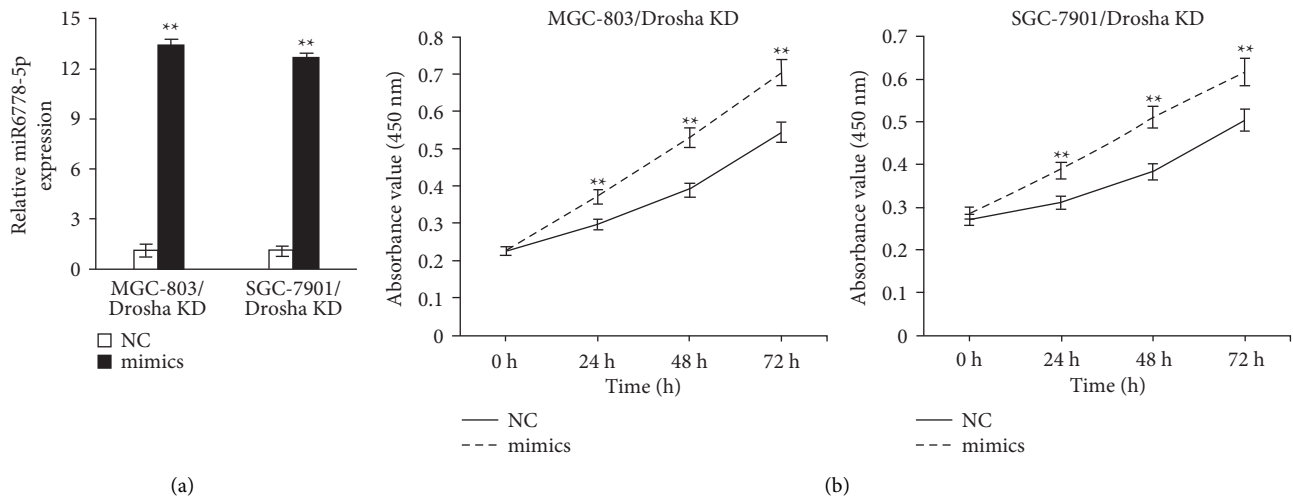
FIGURE 2: Overexpression of miR6778-5p promotes the proliferation of gastric cancer cells with Drosha knockdown. (a) The expression levels of miR6778-5p were evaluated by qRT-PCR (**P<0.05). (b) Growth curves were observed using the CCK8 assay ($^{**}P < 0.05$).

(a)   (b)
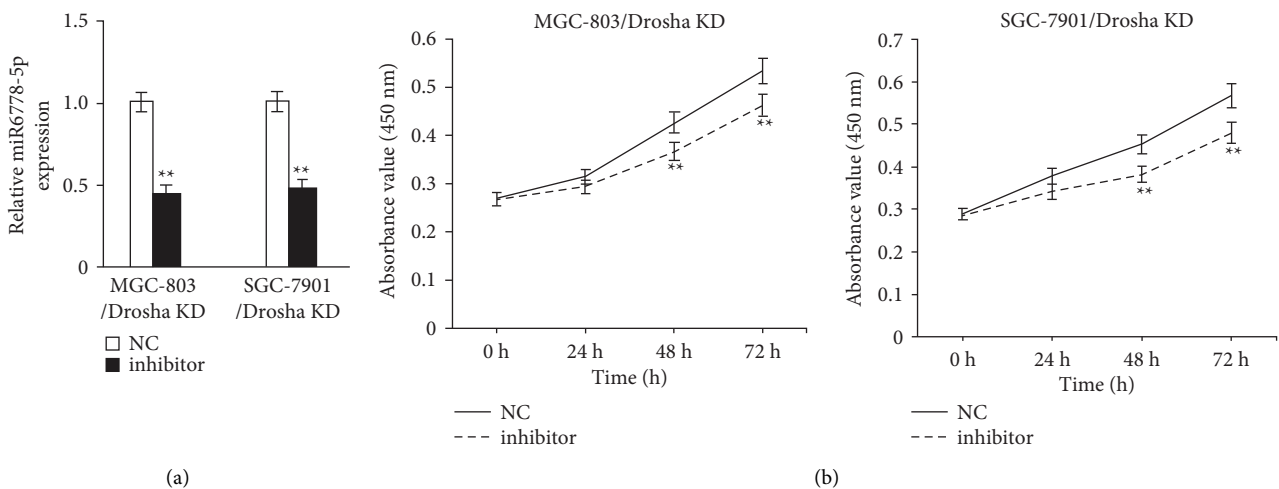
FIGURE 3: miR6778-5p silence inhibits the proliferation of gastric cancer cells with Drosha knockdown. (a) The expression levels of miR6778-5p were evaluated by qRT-PCR (**P<0.05). (b) Growth curves were observed using the CCK8 assay (**P<0.05).

to investigate the intrinsic molecular mechanism of miR6778-5p regulating the multiplication potential of GC cells, we conducted experiments to demonstrate the effect of miR6778-5p on the expression levels of GSK3β, a key negative regulator of cell cycle. According to the experimental results, we observed that after overexpressing miR6778-5p in Drosha low-expressing GC cells, the mRNA and protein expression of GSK3β decreased obviously (Figure 4(b)). On the contrary, after knocking down miR6778-5p in Drosha low-expressing GC cells, the mRNA and protein expression of GSK3β increased apparently (Figure 4(c)). The above results showed that the miR6778-5p/GSK3β axis plays an important role in promoting the proliferation of gastric cancer cells low-expressing Drosha.

4. Discussion

The intrinsic mechanism of gastric cancer carcinogenesis is not fully clarified. Abnormal proliferation, one of the cancer hallmarks, is the chief death cause of gastric cancer patient. Although there is great progress in gastric cancer treatment including radiotherapy, neoadjuvant chemotherapy, molecular targeted therapy, and immunotherapy, gastric cancer, especially advanced gastric cancer, is hard to cure [19]. Therefore, there is an urgent need to reveal the mysterious nature of gastric cancer for finding new treatment targets of GC. In our previous research, we discovered that after Drosha knockdown, the biological behavior of gastric cancer cells will be changed [20]. In this study, Drosha-independent

Position 275-281 of GSK3β 3'UTR　5' ...GAACAUACAAUCUCUCUCCCACG...

|||||||

Hsa-miR-6778-5p　　　　　　　　　　3'　UGGACGGAGGACAGGAGGGUGA

(a)

(b)

(c)
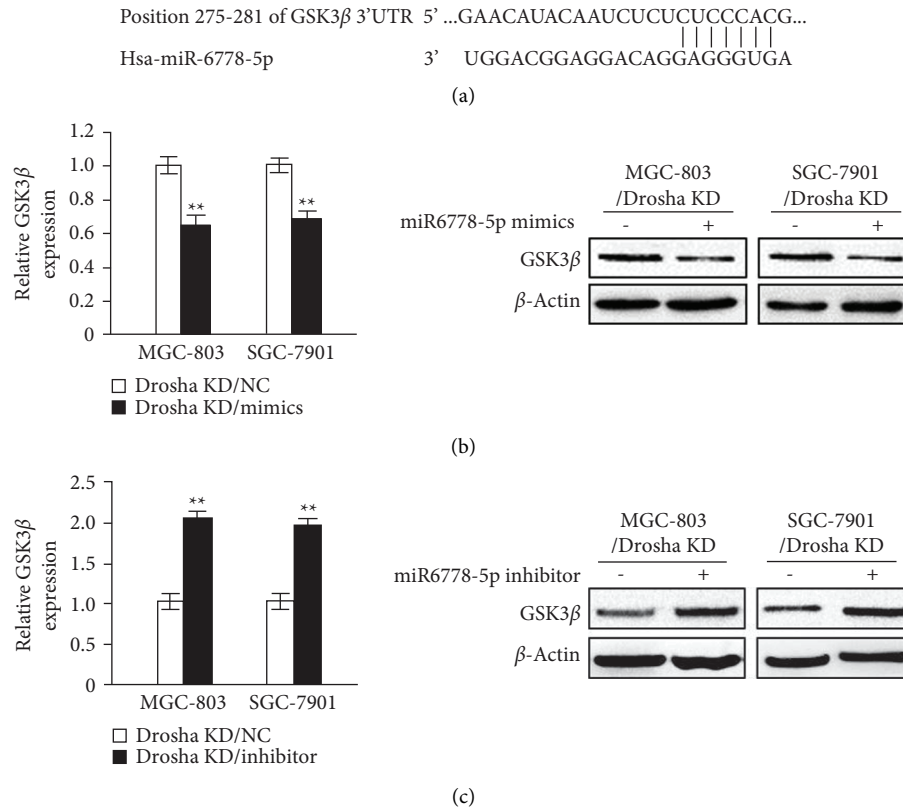
FIGURE 4: The miR6778-5p/GSK3β axis mediates the proliferation of Drosha low-expressing gastric cancer cells. (a) miR6778-5p acts on the 275–281 sites in the UTR of GSK3β mRNA through its seed sequence. (b) The mRNA and protein expression levels of GSK3β were detected by qRT-PCR or WB in MGC-803/Drosha KD or SGC-7901/Drosha KD cell lines after being transfected with miR6778-5p mimics ($^{**}P < 0.05$). Immunoblotting analyses were performed with the indicated antibodies. (c) The mRNA and protein expression levels of GSK3β were determined by qRT-PCR or WB in MGC-803/Drosha KD or SGC-7901/Drosha KD cell lines after being transfected with miR6778-5p inhibitor ($^{**}P < 0.05$). Immunoblotting analyses were performed with the indicated antibodies.

miR6778-5p was found to promote the multiplication of GC cells with Drosha knockdown via inhibiting the expression of its target gene GSK3β. To the best of our knowledge, it is firstly reported that the Drosha-independent miR6778-5p/ GSK3β axis mediates the multiplication of GC cells low-expressing Drosha.

miR6778-5p biogenesis increases in gastric cancer cells with Drosha knockdown. Pri-miRNAs are the primary transcripts which evolve into mature miRNAs after two splicing [21]. Drosha, an endonuclease RNase III enzyme, is necessary for cleaving the primary transcripts of miRNAs, the first step of microRNA maturation process [21]. Besides, Drosha aberrant expression plays a vital part in the pathological process of cancer. Although it is controversial whether Drosha promotes or suppresses the pathological process of cancer [22], Drosha behaves as an oncogene in gastric cancer [16, 20]. Drosha knockdown expectedly reduces the expression of many miRNAs but increases the expression of many other miRNAs [20], indicating that the biogenesis of the upregulated miRNAs is Drosha-independent. This miRNA, obtained by direct splicing through introns and bypassing Drosha, is also called the mirtron pathway [23]. Indeed, the biogenesis of miR-6778-5p is non-canonical in GC cells with Drosha knockdown [16]. However, the non-canonical mechanism of miR-6778-5p

biogenesis remains to be clarified in GC cells with Drosha knockdown.

miR-6778-5p has an active effect upon the multiplication of GC cells with Drosha knockdown via targeting GSK3β. Drosha silence impedes the invasion of GC cells but does not affect other malignant behaviors such as proliferation [16, 20], suggesting that miR-6778-5p could regulate the multiplication of low-expressing Drosha GC cells. MicroRNA is a type of non-coding RNA, which regulates gene expression at the post-transcriptional level by targeting mRNA degradation and/or inhibiting its translation, thereby regulating the development of cancers [24]. It is predicted that GSK3β is a potential target gene of miR6778-5p through searching the microRNA target gene databases (Figure 4(a)). GSK3β is a multifunctional serine/threonine kinase. Now it has been proved that GSK3β acts as a negative regulator in Wnt/β-catenin, PI3K/AKT, and other signaling pathways [25–27]. Upregulation of miRNA-29a can reduce GSK3β protein expression in colorectal cancer cells and inhibit Wnt/β-catenin signaling pathway [28]. In this research, we demonstrated that miR6778-5p inhibits the expression of GSK3β via binding the 3'UTR of GSK3β mRNA in gastric cancer cells with Drosha knockdown. As far as we know, it was firstly found that the miR-6778-5p/GSK3β axis promotes the cell proliferation.

Circular RNAs (circRNAs) are considered a neoteric type of non-coding RNA with organization, structure, time, and space specificity characterized by a covalently closed-loop structure involved in modulating gene expression by regulation of miRNA function, transcription, and protein [29]. Growing study showed that circRNAs contribute to many physiological and pathological processes, including the pathological progression of cancer [30]. CircRNAs perform their biological functions mainly via the circRNA-miRNA-mRNA regulatory network [18, 31, 32]. It remains to be explored whether Drosha-independent miR6778-5p exerts positive effect on the multiplication of GC cells through the circRNA-miRNA-mRNA regulatory network.

5. Conclusions

In a nutshell, this research proves that the Drosha-independent microRNA6778-5p/GSK3β axis mediates the multiplication of GC cells. These discoveries will provide in-depth knowledge of the mechanisms by which Drosha-independent miRNAs promote the abnormal multiplication of GC cells and original therapeutic markers of gastric carcinoma.

Data Availability

The data used in this study are available from the corresponding author upon request.

Disclosure

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Authors' Contributions

M. Ren and S. Tang were responsible for conception and design. W. Wang provided administrative support. M. Ren, L. Xing, and S. Tang were responsible for data analysis and interpretation and provision of study materials or patients. All authors were responsible for collection and assembly of data and manuscript writing and approved the final version of the manuscript.

Acknowledgments

References

[1] S. S. Joshi and B. D. Badgwell, "Current treatment and recent progress in gastric cancer," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 264–279, 2021.

[2] J. Ferlay, M. Colombet, I. Soerjomataram et al., "Cancer statistics for the year 2020: An overview," *International journal of cancer*, vol. 149, 2021.

[3] P. Rawla and A. Barsouk, "Epidemiology of gastric cancer: global trends, risk factors and prevention," *Gastroenterology Review*, vol. 14, no. 1, pp. 26–38, 2019.

[4] D. Lee, J. W. Nam, and C. Shin, "DROSHA targets its own transcript to modulate alternative splicing," *RNA (New York)*, vol. 23, no. 7, pp. 1035–1047, 2017.

[5] S. S. Lee, H. Min, J. Y. Ha, M. S. Choi, and S. Kim, "Dysregulation of the miRNA biogenesis components DICER1, DROSHA, DGCR8 and AGO2 in clear cell renal cell carcinoma in both a Korean cohort and the cancer genome atlas kidney clear cell carcinoma cohort," *Oncology Letters*, vol. 18, no. 4, pp. 4337–4345, 2019.

[6] T. L. Nguyen, T. D. Nguyen, and T. A. Nguyen, "The conserved single-cleavage mechanism of animal DROSHA enzymes," *Communications biology*, vol. 4, no. 1, p. 1332, 2021.

[7] U. Salim, A. Kumar, R. Kulshreshtha, and P. Vivekanandan, "Biogenesis, characterization, and functions of mirtrons," *Wiley interdisciplinary reviews. RNA*, vol. 13, no. 1, Article ID e1680, 2022.

[8] S. Kim, M. L. Song, H. Min et al., "miRNA biogenesis-associated RNase III nucleases Drosha and Dicer are upregulated in colorectal adenocarcinoma," *Oncology Letters*, vol. 14, no. 4, pp. 4379–4383, 2017.

[9] M. Szczyrek, B. Kuźnar-Kamińska, A. Grenda et al., "Diagnostic value of plasma expression of microRNAs complementary to Drosha and Dicer in lung cancer patients," *European Review for Medical and Pharmacological Sciences*, vol. 23, no. 9, pp. 3857–3866, 2019.

[10] Y. K. Kim, B. Kim, and V. N. Kim, "Re-evaluation of the roles of DROSHA, Export in 5, and DICER in microRNA biogenesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 13, pp. E1881–E1889, 2016.

[11] M. Hill and N. Tran, "miRNA interplay: mechanisms and consequences in cancer," *Disease models & mechanisms*, vol. 14, no. 4, Article ID dmm047662, 2021.

[12] T. Debnath, N. C. Deb Nath, E. K. Kim, and K. G. Lee, "Role of phytochemicals in the modulation of miRNA expression in cancer," *Food & Function*, vol. 8, no. 10, pp. 3432–3442, 2017.

[13] Y. Peng, Y. Qin, X. Zhang et al., "MiRNA-20b/SUFU/Wnt axis accelerates gastric cancer cell proliferation, migration and EMT," *Heliyon*, vol. 7, no. 4, Article ID e06695, 2021.

[14] Q. Han, C. Li, Y. Cao et al., "CBX2 is a functional target of miRNA let-7a and acts as a tumor promoter in osteosarcoma," *Cancer Medicine*, vol. 8, no. 8, pp. 3981–3991, 2019.

[15] R. Mancuso, S. Agostini, I. Marventano, A. Hernis, M. Saresella, and M. Clerici, "NCAM1 is the target of miRNA-572: validation in the human oligodendroglial cell line," *Cellular and Molecular Neurobiology*, vol. 38, no. 2, pp. 431–440, 2018.

[16] M. Zhao, Y. Hou, Y. E. Du et al., "Drosha-independent miR-6778-5p strengthens gastric cancer stem cell stemness via regulation of cytosolic one-carbon folate metabolism," *Cancer Letters*, vol. 478, pp. 8–21, 2020.

[17] S. Tang, Y. Hou, H. Zhang et al., "Oxidized ATM promotes abnormal proliferation of breast CAFs through maintaining

intracellular redox homeostasis and activating the PI3K-AKT, MEK-ERK, and Wnt-β-catenin signaling pathways," *Cell Cycle*, vol. 14, no. 12, pp. 1908–1924, 2015.

[18] Z. Z. Liang, C. Guo, M. M. Zou, P. Meng, and T. T. Zhang, "circRNA-miRNA-mRNA regulatory network in human lung cancer: an update," *Cancer Cell International*, vol. 20, no. 1, p. 173, 2020.

[19] Z. Song, Y. Wu, J. Yang, D. Yang, and X. Fang, "Progress in the treatment of advanced gastric cancer," *Tumor Biology*, vol. 39, no. 7, Article ID 101042831771462, 2017.

[20] L. Xu, Y. Hou, G. Tu et al., "Nuclear Drosha enhances cell invasion via an EGFR-ERK1/2-MMP7 signaling pathway induced by dysregulated miRNA-622/197 and their targets LAMC2 and CD82 in gastric cancer," *Cell Death & Disease*, vol. 8, no. 3, Article ID e2642, 2017.

[21] T. Treiber, N. Treiber, and G. Meister, "Regulation of microRNA biogenesis and function," *Thrombosis and Haemostasis*, vol. 107, no. 04, pp. 605–610, 2012.

[22] H. Zhang, Y. Hou, L. Xu et al., "Cytoplasmic drosha is aberrant in precancerous lesions of gastric carcinoma and its loss predicts worse outcome for gastric cancer patients," *Digestive Diseases and Sciences*, vol. 61, no. 4, pp. 1080–1090, 2016.

[23] E. Sakai, Y. Miura, E. Suzuki-Kouyama et al., "A mammalian mirtron miR-1224 promotes tube-formation of human primary endothelial cells by targeting anti-angiogenic factor epsin2," *Scientific Reports*, vol. 7, no. 1, p. 5541, 2017.

[24] T. Y. Mou, R. R. Zhang, and Y. N. Wang, "MiRNA-212 acts as a tumor-suppressor in colorectal carcinoma through targeting SOX4," *European Review for Medical and Pharmacological Sciences*, vol. 23, no. 24, pp. 10751–10760, 2019.

[25] C. Gao, X. Yuan, Z. Jiang et al., "Regulation of AKT phosphorylation by GSK3β and PTEN to control chemoresistance in breast cancer," *Breast Cancer Research and Treatment*, vol. 176, no. 2, pp. 291–301, 2019.

[26] G. V. Vijay, N. Zhao, P. Den Hollander et al., "GSK3β regulates epithelial-mesenchymal transition and cancer stem cell properties in triple-negative breast cancer," *Breast Cancer Research*, vol. 21, no. 1, p. 37, 2019.

[27] B. Xu, J. Xu, N. Cai et al., "Roflumilast prevents ischemic stroke-induced neuronal damage by restricting GSK3β-mediated oxidative stress and IRE1α/TRAF2/JNK pathway," *Free Radical Biology and Medicine*, vol. 163, pp. 281–296, 2021.

[28] X. Han, J. Zheng, Y. Wang, and Z. Gao, "miRNA-29a inhibits colon cancer growth by regulation of the PTEN/Akt/GSK3β and Wnt/β-catenin signaling pathways," *Oncology Letters*, vol. 16, no. 2, pp. 2638–2644, 2018.

[29] S. Sakshi, R. Jayasuriya, K. Ganesan, B. Xu, and K. M. Ramkumar, "Role of circRNA-miRNA-mRNA interaction network in diabetes and its associated complications," *Molecular Therapy - Nucleic Acids*, vol. 26, pp. 1291–1302, 2021.

[30] W. Li, J. Q. Liu, M. Chen, J. Xu, and D. Zhu, "Circular RNA in cancer development and immune regulation," *Journal of Cellular and Molecular Medicine*, vol. 26, no. 6, pp. 1785–1798, 2020.

[31] L. L. Chen, "The expanding regulatory mechanisms and cellular functions of circular RNAs," *Nature Reviews Molecular Cell Biology*, vol. 21, no. 8, pp. 475–490, 2020.

[32] X. Cheng, Q. Chen, and P. Sun, "Natural phytochemicals that affect autophagy in the treatment of oral diseases and infections: a review," *Frontiers in Pharmacology*, vol. 13, Article ID 970596, 2022.

Hindawi

Research Article

Bioinformatic Analysis Identifies of Potential miRNA-mRNA Regulatory Networks Involved in the Pathogenesis of Lung Cancer

Dexun Hao,[1] Yanshuang Li,[2] Jiang Shi,[1] and Junguang Jiang [1]

[1]*Department of Geriatric Respiratory and Sleep, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, Henan, China*
[2]*Department of Anesthesiology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, Henan, China*

Correspondence should be addressed to Junguang Jiang; jiangjunguang@zdyfy.net.cn

Objective. The purpose of the present study was to explore the biomarkers related to lung cancer based on the bioinformatics method, which might be new targets for lung cancer treatment. *Methods*. GSE17681 and GSE18842 were obtained from the Gene Expression Omnibus (GEO) database. The differentially expressed miRNAs (DEMs) and genes (DEGs) in lung cancer samples were screened via the GEO2R online tool. DEMs were submitted to the mirDIP website to predict target genes. Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were conducted via uploading DEGs to the DAVID database. The protein-protein interaction network (PPI) of the DEGs was analyzed by STRING's online tool. Then, the PPI network was visualized using Cytoscape 3.8.0. *Results*. 46 DEMs were identified in GSE17681, and the website predicted that there were 873 target genes of these DEMs. 1029 DEGs were identified in the GSE18842 chip. GO analysis suggested that the co-DEGs participated in the canonical Wnt signaling pathway, regulation of the Wnt signaling pathway, a serine/threonine kinase signaling pathway, the Wnt signaling pathway, and cell-cell signaling by Wnt. KEGG analysis results showed the co-DEGs of GSE17681 and GSE18842 were related to the Hippo signaling pathway and adhesion molecules. In addition, six hub genes that were related to lung cancer were identified as hub genes, including mTOR, NF1, CHD7, ETS1, IL-6, and COL1A1. *Conclusions*. The present study identified six hub genes that were related to lung cancer, including mTOR, NF1, CHD7, ETS1, IL-6, and COL1A1, which might be a potential target for lung cancer.

1. Introduction

At present, the incidence rate and mortality of lung cancer rank first in China [1], while non-small cell lung cancer (NSCLC) is the most common type of lung cancer, accounting for about 85% of them [2]. Adenocarcinoma and squamous cell carcinoma are common pathological types of NSCLC [3]. Because there are no obvious clinical symptoms in the early stage, the vast majority of NSCLC patients have entered the late stage at the time of diagnosis, and the 5-year survival rate is about 15%. Radiotherapy, chemotherapy, and surgery play a key role in the treatment of NSCLC, but 50% of lung cancer patients die of tumor recurrence [4]. Therefore, it is very important to find biomarkers that can accurately predict the prognosis of patients. Elucidating the

genetic changes related to the occurrence and development of the disease at the molecular level will be conducive to the diagnosis, treatment, and prognosis of the disease. At present, the pathogenesis of NSCLC is still not clear, so mining the genetic changes related to NSCLC from the genome level will likely provide more molecular markers for its diagnosis and prognosis.

MiRNA is a small noncoding RNA that can regulate gene expression by targeting mRNA. It is closely related to cell proliferation, differentiation, migration, and invasion and is involved in tumor development. The development of high-throughput technology provides abundant gene expression profile data for the study of the pathogenesis of NSCLC [5]. Many scholars have conducted in-depth studies on the pathogenesis of NSCLC at the gene level and found that
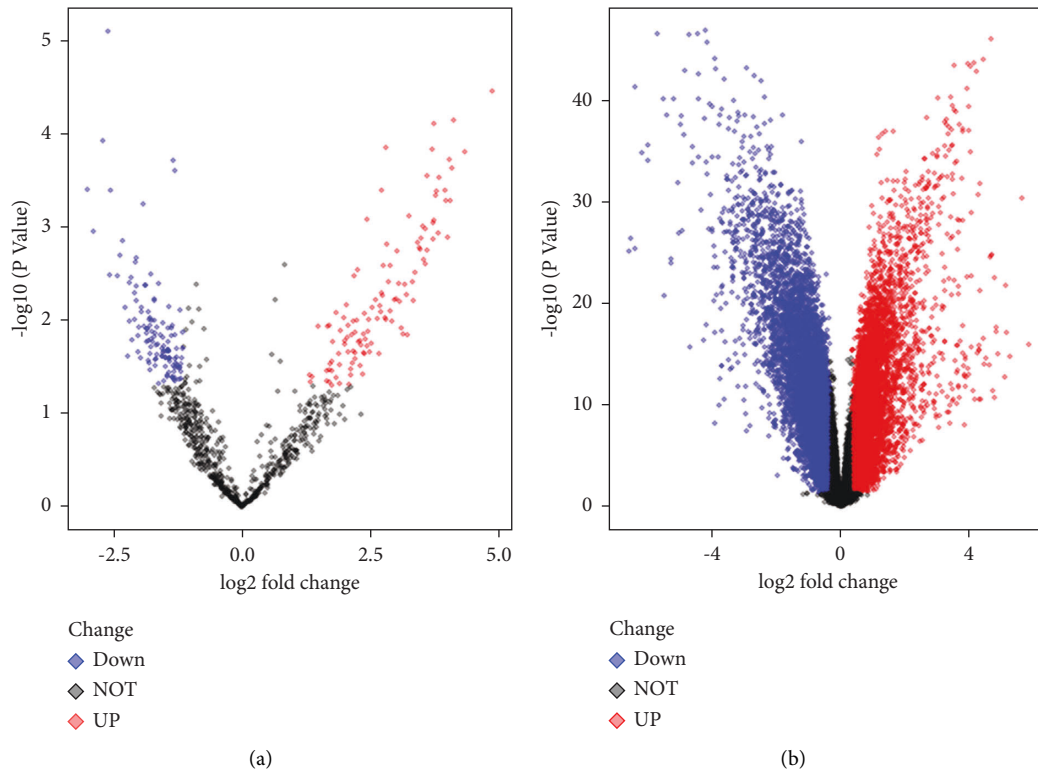
FIGURE 1: Volcano plots of the DEGs in GSE17681 (a) and GSE18842 (b).

miRNA is related to the occurrence and progression of NSCLC [6, 7]. In recent years, more and more researchers have begun to study NSCLC-targeted new drug development. However, most patients will develop resistance to targeted drugs, resulting in poor efficacy for targeted drugs [8]. Therefore, finding new therapeutic targets for NSCLC is of great significance for improving the survival rate of NSCLC patients. Although there are reports on the screening of the DEGs in NSCLC, the results are also different due to the inconsistent number of samples in each study and the influence of confounding factors such as patient source, tumor stage, and grade.

Therefore, this study screened two NSCLC chip datasets (GSE17681 and GSE18842) to find common DEGs, aiming to provide data support for the precise treatment and prognosis of NSCLC.

2. Material and Methods

The microarray data from Gene Expression Omnibus (GEO) databases, GSE17681 and GSE18842 were utilized in this study. The dataset GSE17681, which was the miRNA profiling and based on the GPL570 platform, collected 17 NSCLC specimens and 19 normal tissues. The dataset GSE18842 collected 46 NSCLC specimens and 45 normal tissues, which were based on the GPL570 platform.

2.1. Identification of DEGs.
Firstly, the GEO2R online tool was utilized to screen the differentially expressed miRNAs (DEMs) and genes between lung cancer samples and healthy samples. The "GEOquery" and "limma" packages in R

software were utilized to analyze data tables. adj. $P < 0.05$ and $|logFC| > 2$ were set as the criteria for significant differences. DEMs were submitted to the mirDIP website (https://ophid.utoronto.ca/mirDIP/index.jsp#r) to predict target genes. The intersection of the DEGs in GSE17681 and GSE18842 was conducted to screen out the common DEGs. The heatmap based on top 10 and bottom 10 DEGs were created by the "pheatmap" package. A volcano map was generated using the "ggplot2" package.

2.2. Functional and Pathway Enrichment Analyses.
The biological information annotation database David (https://david.ncifcrf.gov/) was used for Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, including cell components, molecular functions, biological processes, and signal pathways. The "pheatmap" package of R software was used to visualize the enrichment analysis results.

2.3. Protein-Protein Interaction Network (PPI) Network Construction.
The PPI network of the DEGs was analyzed using STRING (https://cn.string-db.org/). Then, the visualization of the PPI network was performed by Cytoscape 3.8.0, and the genes with a high level of connectivity with surrounding genes were selected as hub genes.

3. Results

3.1. Identification of DEGs.
The GSE17681 chip yielded 46 DEMs, of which 37 were upregulated and 9 were
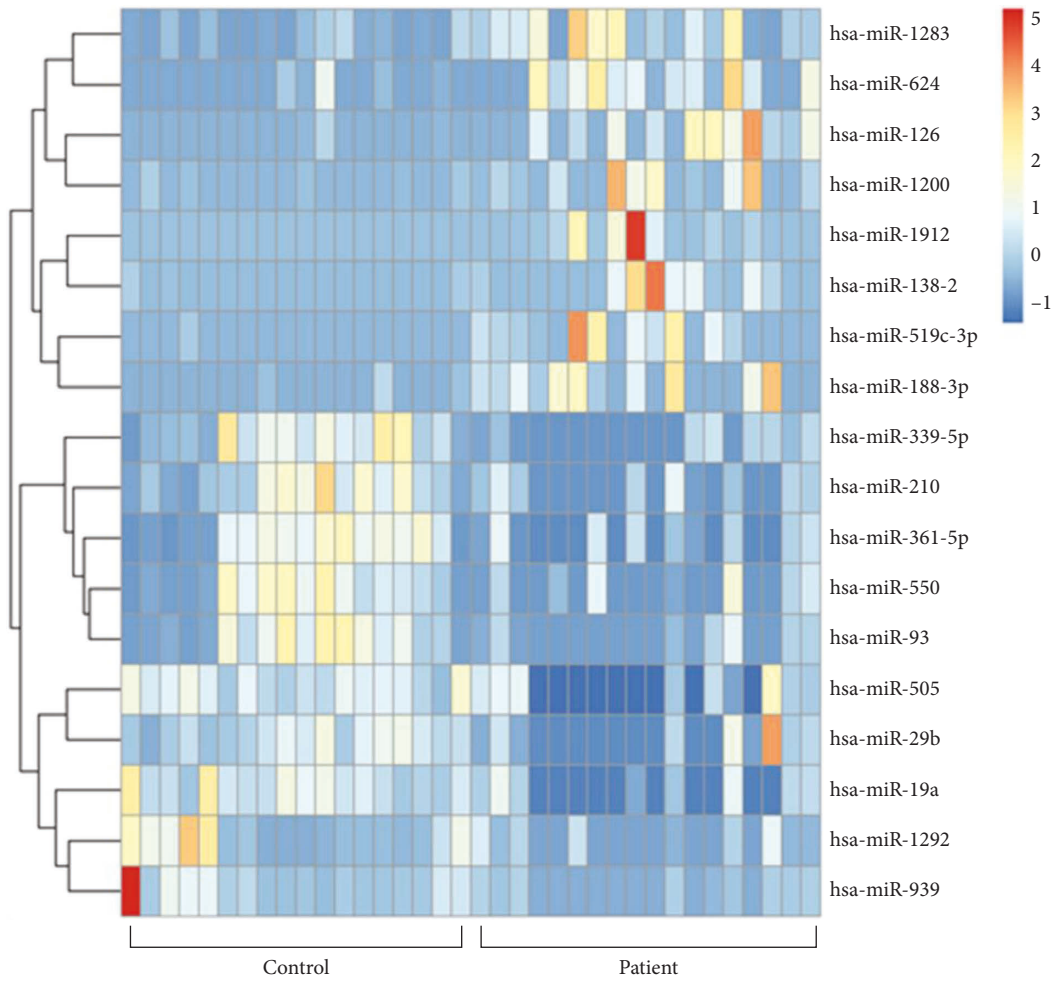
(a)

FIGURE 2: Continued.
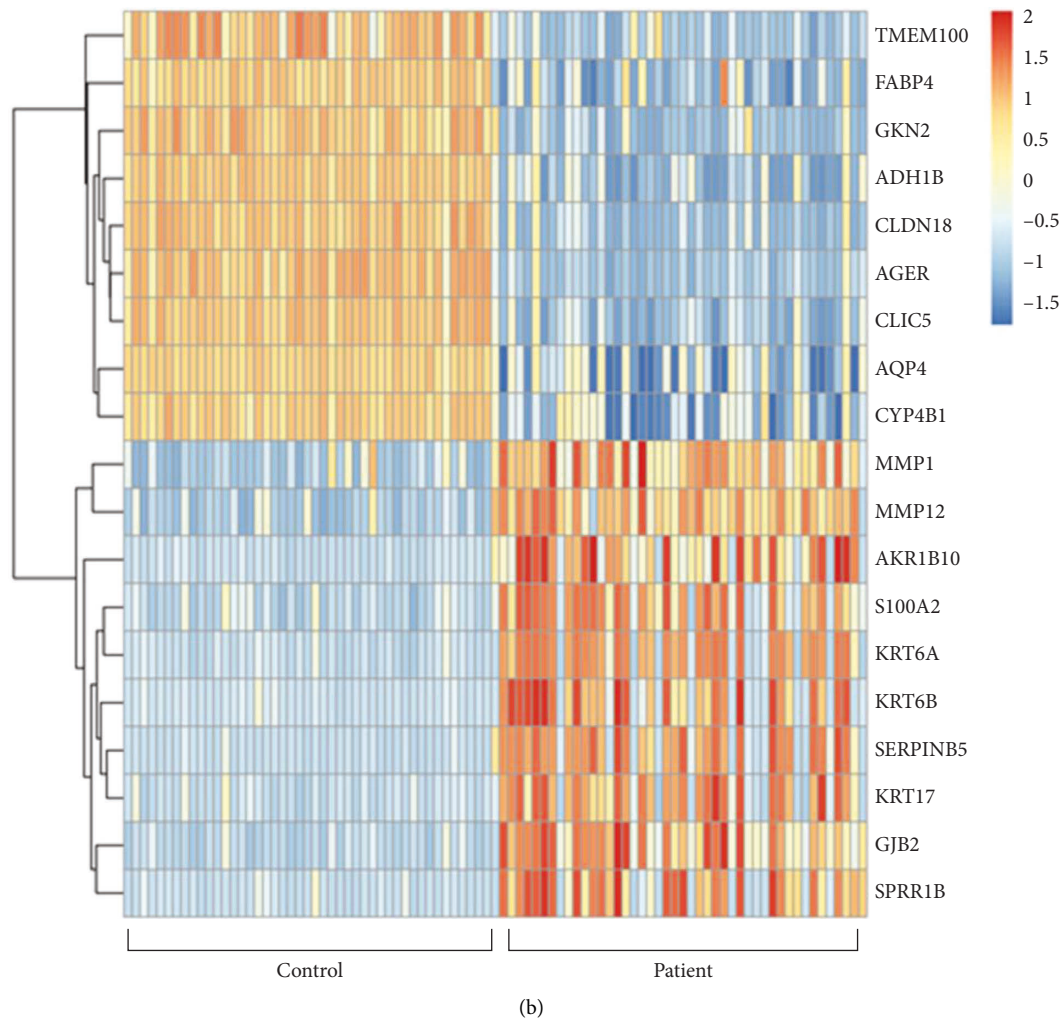
Figure 2: The heatmap plots of the top 10 genes. (a) GSE17681. (b) GSE18842.

downregulated (Figure 1(a)). The website predicted that there were 873 target genes for these DEMs. 1029 DEGs were obtained on the GSE18842 chip, of which 419 were upregulated and 610 were downregulated (Figure 1(b)). Further analysis showed that 84 common genes (35 upregulated genes and 49 downregulated genes), which were named co-DEGs, were differentially expressed in lung cancer in these two-chip data. The top 10 and bottom 10 DEGs expressions were visualized by a heatmap (Figure 2).

3.2. GO Annotation Analyses of DEGs. The DEGs of GSE17681 participated in the regulation of mRNA metabolic process, eye development, visual system development, myeloid cell differentiation, sensory system development, regulation of mRNA stability, mRNA catabolic process, and regulation of mRNA splicing via spliceosome (Figure 3(a)). The DEGs of GSE18842 participated in mitotic sister chromatid segregation, sister chromatid segregation, and nuclear chromosome segregation (Figure 3(b)). The co-DEGs participated in the canonical Wnt signaling pathway,

regulation of the canonical Wnt signaling pathway, regulation of the Wnt signaling pathway, the transmembrane receptor protein serine/threonine kinase signaling pathway, the Wnt signaling pathway, and cell-cell signaling by Wnt (Figure 3(c)).

3.3. KEGG Pathway Enrichment Analyses of DEGs. The DEGs of GSE17681 were related to the Wnt signaling pathway, hepatocellular carcinoma, acute myeloid leukemia, regulation of actin cytoskeleton, cellular senescence, pathways in cancer, the MAPK signaling pathway, gastric cancer, colorectal cancer, and breast cancer (Figure 4(a)). The DEGs of GSE18842 were related to protein digestion and absorption, progesterone-mediated oocyte maturation, PPAR signaling pathway, African trypanosomiasis, p53 signaling pathway, oocyte meiosis, malaria, lipid and atherosclerosis, IL-17 signaling pathway, estrogen signaling pathway, cellular senescence, cell cycle, arachidonic acid metabolism, amoebiasis, and the AGE-RAGE signaling pathway in diabetic complications (Figure 4(b)). The co-DEGs of
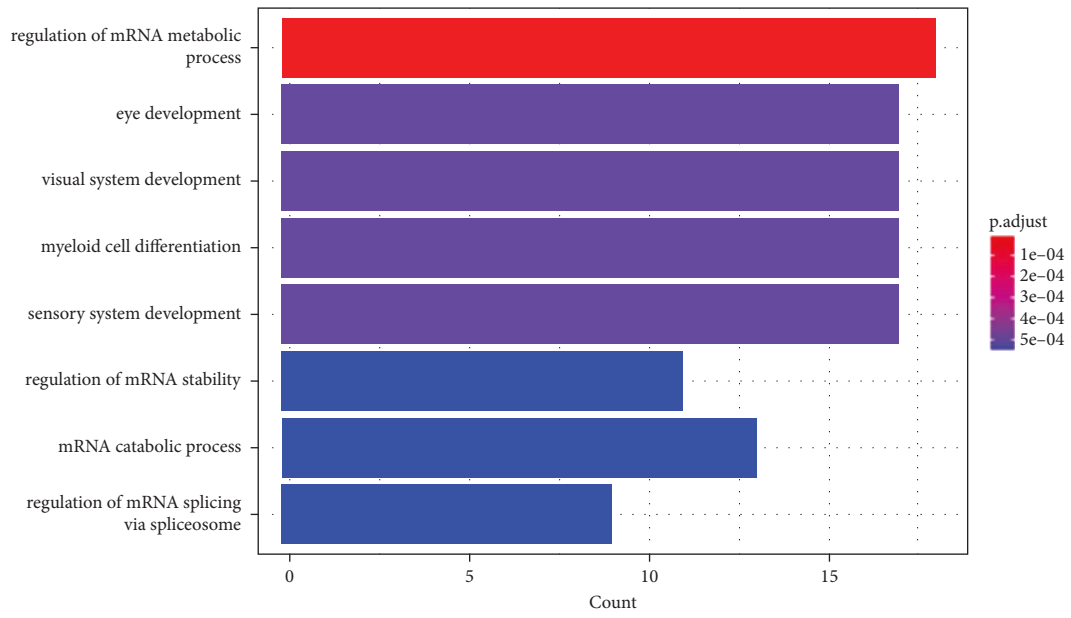
(a)
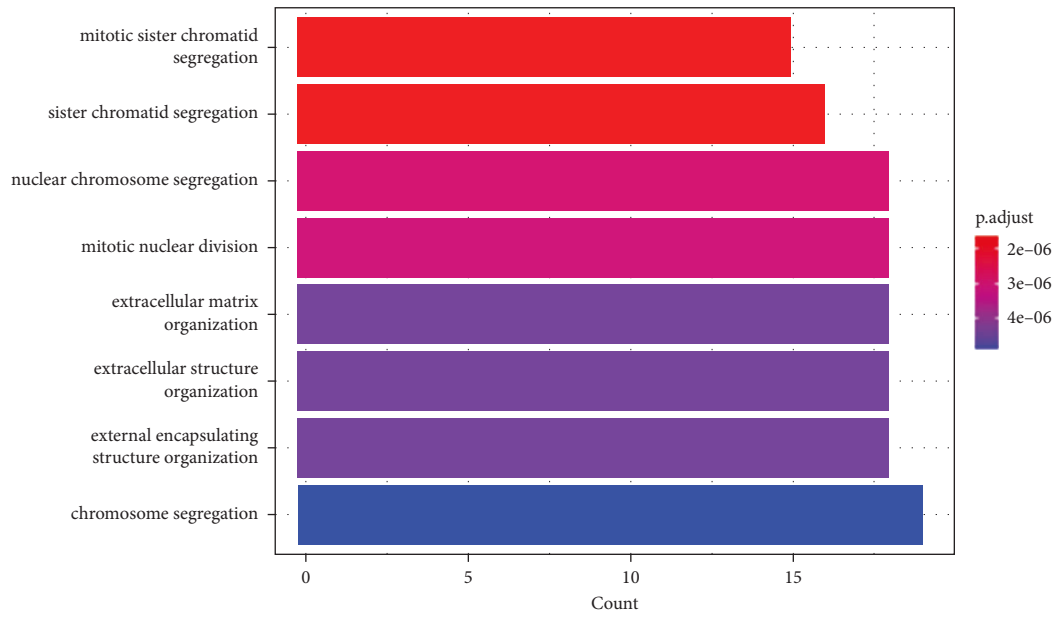
(b)

FIGURE 3: Continued.

(c)
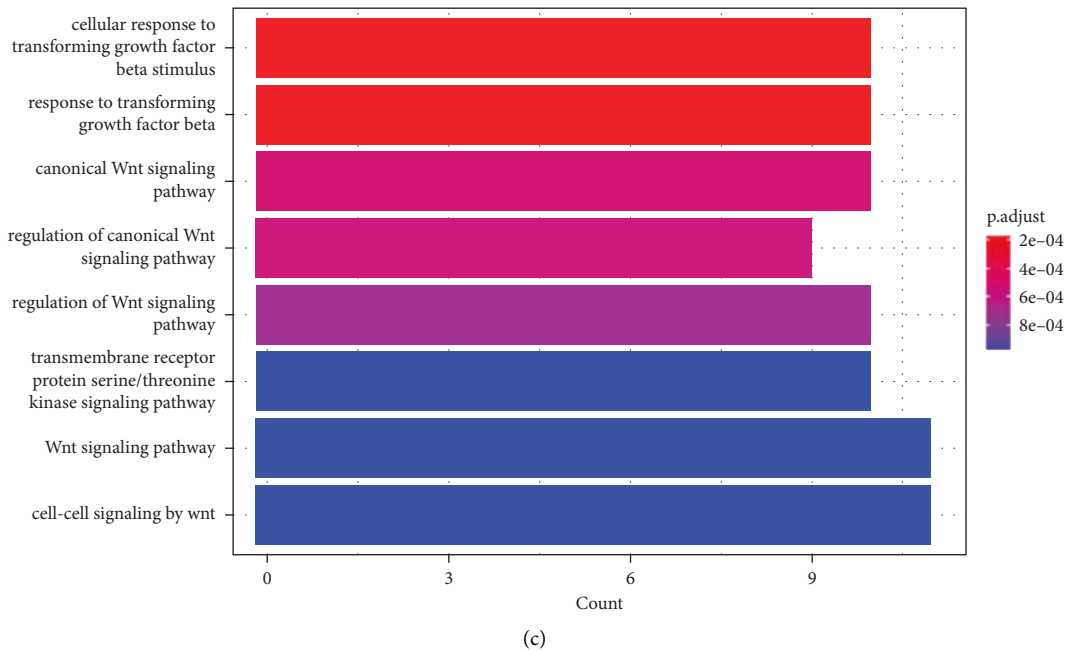
FIGURE 3: Gene Ontology (GO) analyses of the DEGs of GSE17681 (a), GSE18842 (b), and the co-DEGs (c).

GSE17681 and GSE18842 were related to the Hippo signaling pathway and adhesion molecules (Figure 4(c)).

3.4. PPI Network and Hub Genes. To identify the hub genes relevant to lung cancer, DEGs were uploaded to the STRING website to create a PPI network, and the visualization results are shown in Figure 5. Six hub genes were screened in relation to lung cancer, including mTOR, NF1, CHD7, ETS1, IL-6, and COL1A1.

4. Discussion

Over the years, although NSCLC treatment has been continuously improved, its 5-year survival rate has remained very low, at only 16% [9], mainly because its pathogenesis remains unclear. Lung cancer is a complex disease, and its occurrence may be a combination of polygenic and multipathway roles [10]. Therefore, finding key genes for lung carcinogenesis from the perspective of tumor genome-wide alterations may be an effective way to study the pathogenesis of lung cancer. With the rapid development of biological information technology, the application of high-throughput technologies such as microarray and whole genome sequencing to mine key genes in the process of disease occurrence and progression has brought new methods for exploring the molecular pathogenesis of the disease, improving clinical diagnosis, and targeting therapy [11]. At present, there are few studies on the pathogenesis and related molecular markers of NSCLC at home and abroad, which limits the timely diagnosis and treatment of NSCLC in clinical practice.

In this study, 46 DEMs (37 upregulated and 9 downregulated) were screened on the GSE17681 chip. The website

predicted that there were 873 target genes for these DEMs. 1029 DEGs (419 upregulated and 610 downregulated) were screened on the GSE18842 chip. GO analysis results showed that the co-DEGs participated in the canonical Wnt signaling pathway, regulation of the canonical Wnt signaling pathway, the transmembrane receptor protein serine/threonine kinase signaling pathway, the Wnt signaling pathway, and cell-cell signaling by Wnt. KEGG analysis results showed the co-DEGs of GSE17681 and GSE18842 were related to the Hippo signaling pathway and adhesion molecules. Further verification found that six hub genes were screened in relation to lung cancer, including mTOR, NF1, CHD7, ETS1, IL-6, and COL1A1.

The activation of the mTOR signaling pathway is involved in the occurrence and development of human tumors and promotes the occurrence of tumors through a variety of mechanisms. The activation of various components in the mTOR signaling pathway is also a poor prognostic factor for many tumors. Inhibition of the mTOR signaling pathway can reverse drug resistance and improve the effect of chemotherapy and radiotherapy in vivo and in vitro [12]. MTOR was upregulated in numerous tumors, such as ovarian cancer, breast cancer, lung cancer, and so on [13]. The MTOR gene showed genome amplification in lung cancer and preinvasion bronchopathy, suggesting that the mTOR pathway was associated with the development of lung cancer [14]. The protein encoded by NF1 is associated with cell growth and differentiation. The NF1 gene is a tumor suppressor gene, and its encoded neurofibroma protein is a Ras GTP enzyme activator protein (RAS gap) [15]. Neurofibromatosis protein is functionally and structurally homologous to p120RasGAP. Gap proteins can terminate the Ras-mediated signal transduction pathway by activating the Ras-activated GTP binding form into an inactive GDP

(a)

(b)

(c)
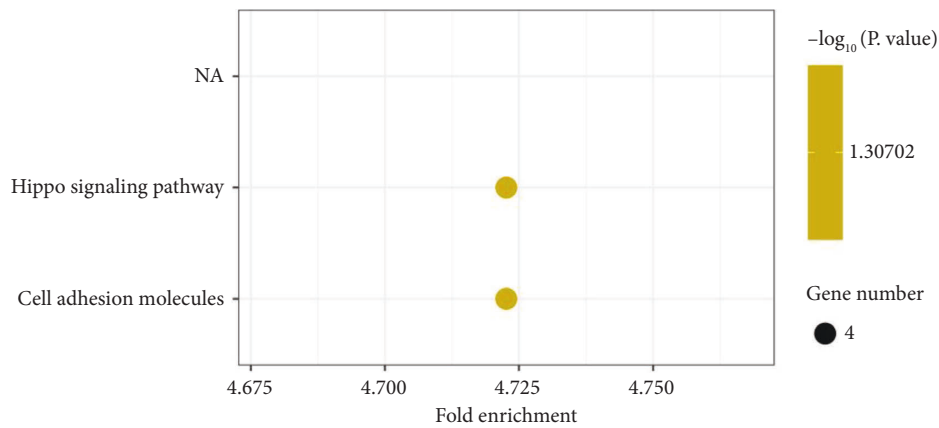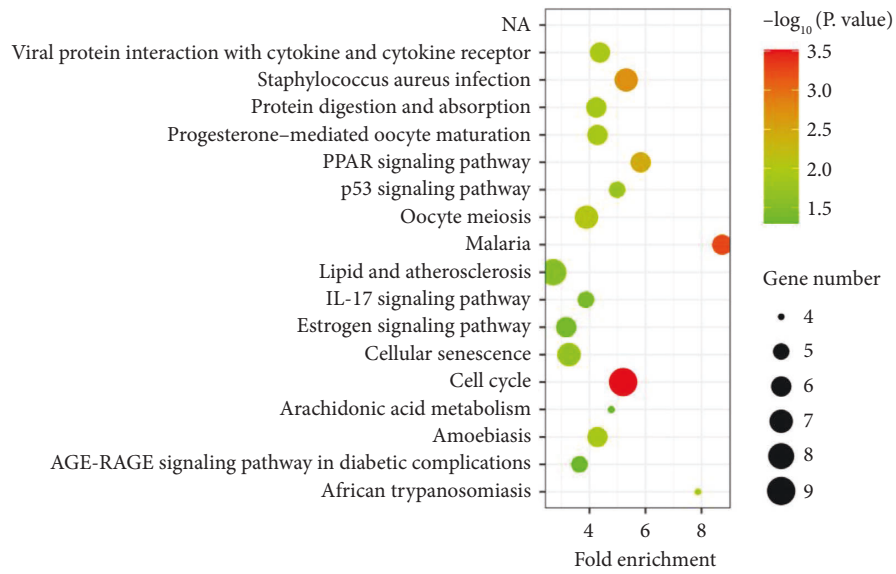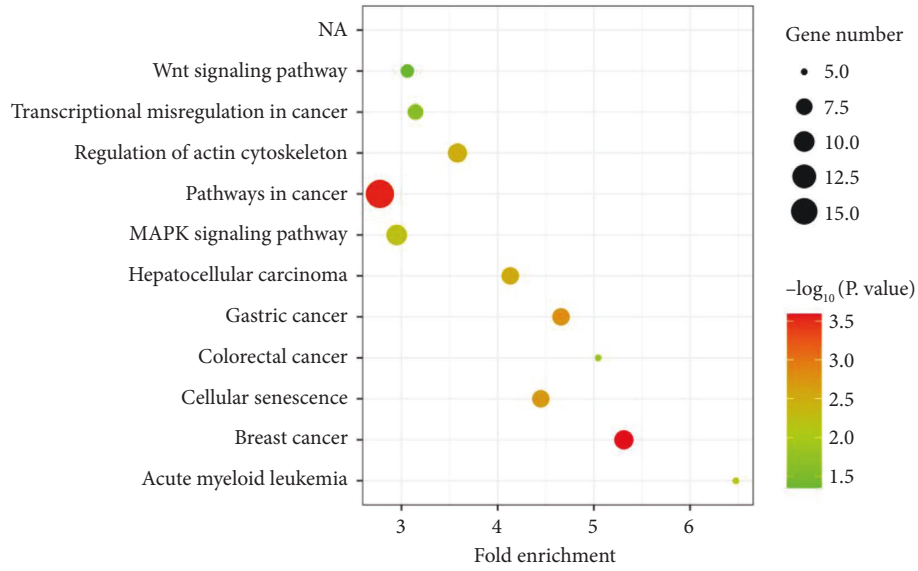
Figure 4: KEGG analyses of the DEGs of GSE17681 (a), GSE18842 (b), and the co-DEGs (c).
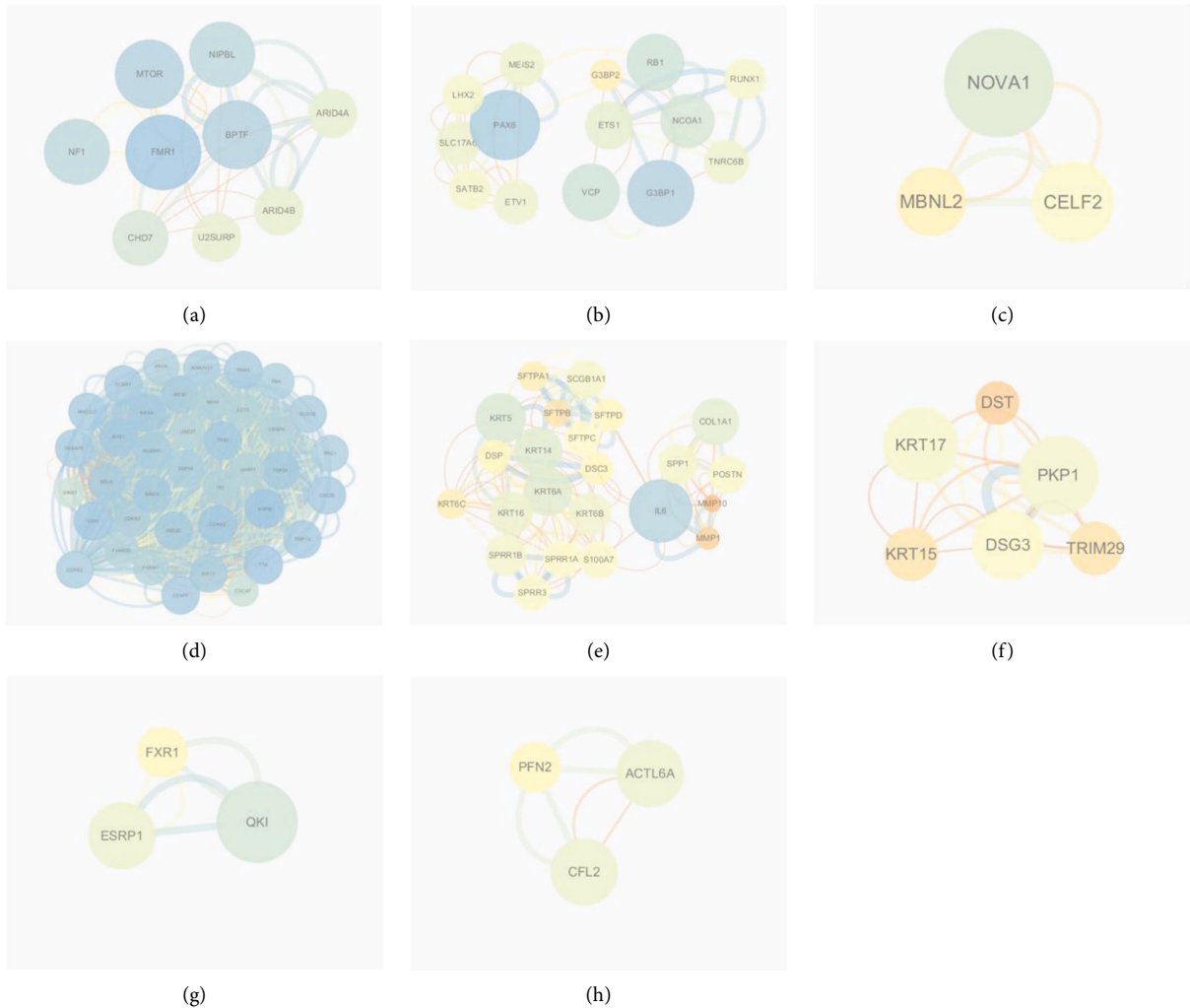
Figure 5: PPI networks are based on the screened DEGs. (a, b, c). The PPI networks are based on the DEGs of GSE17681 screened above. (d, e, f) The PPI networks are based on the DEGs of GSE18842 screened above. (g, h, i). The PPI networks are based on the co-DEGs of GSE17681 and GSE18842 screened above.

binding form. Therefore, the tumor suppressive function of NF1 is considered to be mainly dependent on its down-regulation of the proto-oncogene RAS.

CHD7 is a member of the chromatin helicase DNA binding protein family. Its biological functions are mostly related to human congenital malformations [16], and it is abnormally expressed and activated in a variety of tumors such as SCLC and pancreatic cancer, regulating tumor cell proliferation, invasion, and other functions [17, 18]. Pleasance et al. found that CHD7 rearranged in NSCLC, which was relevant to abnormal cell damage repair, and played a certain role in promoting the progress of lung cancer [17]. ETS-1 is a proto-oncogene, which is positive in many tumors [19]. The ETS-1 gene is involved in cell growth and extracellular matrix invasion, which can promote tumor invasion and metastasis [20]. ETS-1 coding products constitute a large family of transcription regulators that participate in cell migration and apoptosis, angiogenesis, and organogenesis. ETS-1 is generally highly expressed in many malignant tumors. ETS-1 can also regulate some ECM target

genes, including matrix proteins and other cell components involved in cell-matrix reactions [21]. When the expression of endogenous ETS-1 is inhibited, the expression of matrix-related proteins MMP-1 and MMP-2 is downregulated at the RNA level.

IL-6 is also a downstream effector of the oncogene RAS, which affects the immune regulation and hematopoietic regulation of the body [22] and promotes the differentiation and proliferation of T and B lymphocytes by activating target genes, so as to enhance the activity of neutrophils and monocytes and enhance the inflammatory response of local tissues. COL1A1 can form collagen fibers; as an effective component of bone marrow, it is also involved in the proliferation, metastasis, and angiogenesis of a variety of tumor cells [23]. Type I collagen gene deficiency can promote the metastasis of breast cancer cells. Want3a/β-Mrtf-a silencing in the catenin pathway can reduce the binding of acetylated histone H3K9 and RNA polymerase on the COL1A1 promoter and reduce the expression of COL1A1, while the transcription factor Osterix can directly bind to the

COL1A1 promoter and upregulate the expression of COL1A1 [24]. COL1A1 was reported to be involved in the differentiation and metastasis of human bladder cancer [25]. Liu et al. reported that COL1A1 can mediate the metastasis of breast cancer, which might become the potential target for breast cancer [23].

In conclusion, the present study screened 46 DEMs in GSE17681 and 1029 DEGs in GSE18842 and identified six hub genes related to lung cancer, including mTOR, NF1, CHD7, ETS1, IL-6, and COL1A1, which might be potential targets for lung cancer. This study used bioinformatics to find the key to lung cancer, provided important potential targets for early diagnosis and prognosis of lung cancer, and formulated new diagnosis and treatment strategies for patients. However, the limitation of this study lies in the lack of experimental evidence for the key genes screened. The next line of research can continue to explore the impact of these key genes on lung cancer in vivo and in vitro experiments.

Data Availability

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

Conflicts of Interest

The authors declare that they have conflicts of interest.

References

[1] Q. Zhen, L. Gao, R. F. Wang et al., "LncRNA DANC promotes lung cancer by sequestering miR-216a," *Cancer Control*, vol. 25, no. 1, 2018.

[2] B. C. Bade and C. S. Dela Cruz, "Lung cancer 2020: epidemiology, etiology, and prevention," *Clinics in Chest Medicine*, vol. 41, no. 1, pp. 1–24, 2020.

[3] F. Wu, L. Wang, and C. Zhou, "Lung cancer in China: current and prospect," *Current Opinion in Oncology*, vol. 33, no. 1, pp. 40–46, 2021.

[4] R. S. Herbst, D. Morgensztern, and C. Boshoff, "The biology and management of non-small cell lung cancer," *Nature*, vol. 553, no. 7689, pp. 446–454, 2018.

[5] J. Zhao, C. Guo, Z. Ma, H. Liu, C. Yang, and S. Li, "Identification of a novel gene expression signature associated with overall survival in patients with lung adenocarcinoma: a comprehensive analysis based on TCGA and GEO databases," *Lung Cancer*, vol. 149, pp. 90–96, 2020.

[6] R. Zhang and A. Ma, "High expression of MYEOV reflects poor prognosis in non-small cell lung cancer," *Gene*, vol. 770, Article ID 145337, 2021.

[7] W. Zhang, Q. Zhang, L. Che et al., "Using biological information to analyze potential miRNA-mRNA regulatory networks in the plasma of patients with non-small cell lung cancer," *BMC Cancer*, vol. 22, no. 1, p. 299, 2022.

[8] S. G. Wu and J. Y. Shih, "Management of acquired resistance to EGFR TKI-targeted therapy in advanced non-small cell lung cancer," *Molecular Cancer*, vol. 17, no. 1, p. 38, 2018.

[9] R. S. Herbst, D. Prager, R. Hermann et al., "TRIBUTE: a phase III trial of erlotinib hydrochloride (OSI-774) combined with carboplatin and paclitaxel chemotherapy in advanced non-small-cell lung cancer," *Journal of Clinical Oncology*, vol. 23, no. 25, pp. 5892–5899, 2005.

[10] J. Rodriguez-Canales, E. Parra-Cuentas, and I. I. Wistuba, "Diagnosis and molecular classification of lung cancer," *Cancer Treatment and Research*, vol. 170, pp. 25–46, 2016.

[11] Y. Mao, P. Xue, L. Li et al., "Bioinformatics analysis of mRNA and miRNA microarray to identify the key miRNA-gene pairs in small-cell lung cancer," *Molecular Medicine Reports*, vol. 20, no. 3, pp. 2199–2208, 2019.

[12] M. Wangpaichitr, C. Wu, M. You et al., "Inhibition of mTOR restores cisplatin sensitivity through down-regulation of growth and anti-apoptotic proteins," *European Journal of Pharmacology*, vol. 591, no. 1-3, pp. 124–127, 2008.

[13] E. Conde, B. Angulo, M. Tang et al., "Molecular context of the EGFR mutations: evidence for the activation of mTOR/S6K signaling," *Clinical Cancer Research*, vol. 12, no. 3, pp. 710–717, 2006.

[14] H. Dudek, S. R. Datta, T. F. Franke et al., "Regulation of neuronal survival by the serine-threonine protein kinase Akt," *Science*, vol. 275, no. 5300, pp. 661–665, 1997.

[15] K. J. Radomska, F. Coulpier, A. Gresset et al., "Cellular origin, tumor progression, and pathogenic mechanisms of cutaneous neurofibromas revealed by mice with Nf1 knockout in boundary cap cells," *Cancer Discovery*, vol. 9, no. 1, pp. 130–147, 2019.

[16] T. Zhen, E. M. Kwon, L. Zhao et al., "Chd7 deficiency delays leukemogenesis in mice induced by Cbfb-MYH11," *Blood*, vol. 130, no. 22, pp. 2431–2442, 2017.

[17] E. D. Pleasance, P. J. Stephens, S. O'Meara et al., "A small-cell lung cancer genome with complex signatures of tobacco exposure," *Nature*, vol. 463, no. 7278, pp. 184–190, 2010.

[18] L. E. Colbert, A. V. Petrova, S. B. Fisher et al., "CHD7 expression predicts survival outcomes in patients with resected pancreatic cancer," *Cancer Research*, vol. 74, no. 10, pp. 2677–2687, 2014.

[19] S. U. Nazir, R. Kumar, A. Singh et al., "Breast cancer invasion and progression by MMP-9 through Ets-1 transcription factor," *Gene*, vol. 711, Article ID 143952, 2019.

[20] A. Furlan, C. Vercamer, L. Heliot, N. Wernert, X. Desbiens, and A. Pourtier, "Ets-1 drives breast cancer cell angiogenic potential and interactions between breast cancer and endothelial cells," *International Journal of Oncology*, vol. 54, no. 1, pp. 29–40, 2019.

[21] A. F. Okuducu, U. Zils, S. A. M. Michaelis, S. Michaelides, and A. von Deimling, "Ets-1 is up-regulated together with its target gene products matrix metalloproteinase-2 and matrix metalloproteinase-9 in atypical and anaplastic meningiomas," *Histopathology*, vol. 48, no. 7, pp. 836–845, 2006.

[22] S. Kaur, Y. Bansal, R. Kumar, and G. Bansal, "A panoramic review of IL-6: structure, pathophysiological roles and inhibitors," *Bioorganic & Medicinal Chemistry*, vol. 28, no. 5, Article ID 115327, 2020.

[23] J. Liu, J. X. Shen, H. T. Wu et al., "Collagen 1A1 (COL1A1) promotes metastasis of breast cancer and is a potential therapeutic target," *Discovery Medicine*, vol. 25, no. 139, pp. 211–223, 2018.

[24] Q. Geng, Z. Shen, L. Li, and J. Zhao, "COL1A1 is a prognostic biomarker and correlated with immune infiltrates in lung cancer," *PeerJ*, vol. 9, Article ID e11145, 2021.

[25] K. Mori, H. Enokida, I. Kagara et al., "CpG hypermethylation of collagen type I alpha 2 contributes to proliferation and migration activity of human bladder cancer," *International Journal of Oncology*, vol. 34, no. 6, pp. 1593–1602, 2009.

Research Article

Recurrent Neural Model to Analyze the Effect of Physical Training and Treatment in Relation to Sports Injuries

Jyoti A. Dhanke,[1] **Rajesh Kumar Maurya,**[2] **S. Navaneethan,**[3] **Dinesh Mavaluru** ⓘ **,**[4] **Shibili Nuhmani** ⓘ **,**[5] **Nilamadhab Mishra** ⓘ **,**[6] **and Ellappan Venugopal** ⓘ[7]

[1]*Department of Science (Mathematics), Bharati Vidyapeeth's College of Engineering, Lavale, Pune 412115, India*
[2]*Department of Computer Applications, ABES Engineering College, Ghaziabad 201009, Uttar Pradesh, India*
[3]*Department of Electronics and Communication Engineering, Saveetha Engineering College, Chennai, Tamil Nadu, India*
[4]*Department of Information Technology, College of Computing and Informatics, Saudi Electronic University, Riyadh, Saudi Arabia*
[5]*Department of Physical Therapy, Imam Abdulrahman Bin Faisal University, Dammam 34212, Saudi Arabia*
[6]*School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh 466114, India*
[7]*Department of Electronics and Communication Engineering, School of Electrical Engineering and Computing, Adama Science and Technology University, Adama, Ethiopia*

Correspondence should be addressed to Ellappan Venugopal; ellappan.venugopal@astu.edu.et

Artificial intelligence has rapidly grown and has made the scenario that no field can function without it. Like every field, it also plays a vital role in the sports field nowadays. In certain sports, injuries happen very often due to heavy training and sudden speedy actions, especially in athletics and football. Here arises a need to analyze the effect of physical training in sportsperson by collecting data from their daily training. With the help of artificial intelligence, a recurrent neural model is developed to analyze the effect of physical training and treatment concerning sports injury. A Recurrent Neural Network (RNN) can be a subsection of Artificial Neural Networks (ANN) that uses the neural nodes connected in a temporal sequence. The temporal sequence is one of the essential terms in this research, which denotes a data sequence of events in a given timeframe. The recurrent neural model is an intelligent machine learning method that comprises a neural schema replicating humans. This neural schema studies the data it collects from the athletes/players and processes it by analyzing previous injuries. Sports injuries have to be analyzed because, in some cases, it becomes more dangerous to the sportsperson that they may even lose their career due to disability. Sometimes it may cause a massive loss to the club or company that hired the sportsperson for the sport. The prediction process can give the player rest until he recovers, thus becoming the safest approach in sports. Therefore, it is essential to analyze the sportsperson's track data to keep an eye on his health. In this research, RNN model is compared with the existing Support Vector Machine (SVM) in concerning to the effect of physical training and treatment for sports. The results show that the proposed model has achieved 99% accuracy, which is higher than the existing algorithm.

1. Introduction

A Recurrent Neural Network (RNN) model is a type of Artificial Neural Network (ANN) that aids in the system's dynamic behavior for a shorter duration. Additionally, the nodes in the networks are connected, which can construct a graph of directed or undirected categories over time.

Athletes compete in a variety of events that require extreme fitness and stamina. Illness and aging produce anatomical and functional changes in the human body, putting senior people at risk of musculoskeletal and cardiovascular overload [1]. A substantial percentage of injuries (both acute and exceptional) are minor and can be treated with a bit of break from training and competition. The injuries of the

sportsperson have to be treated as soon as possible and as effectively as possible [2]. However, it is important to note that inactivity and immobility have a more significant influence on structures and functions in the old than in the younger. Most physically active seniors are selected because they have better health and physical capacity than sedentary individuals their age, increasing their physical capacity. They will, however, be impacted by some of the drawbacks of physical overloading, mainly because the aging body's capacity to respond to high levels of loading is impaired [3]. Acute injuries are prevalent in senior persons who participate in ball games, downhill skiing, and gymnastics, which involve a lot of coordination, reaction time, and balance. Muscle has been described as the most usually acutely damaged tissue among active senior athletes. The most prevalent type of injury is to the lower extremities [4].

2. Literature Review

The RNN approach is designed to capture temporal correlations and is more successful for sequential data than traditional neural networks (NNs). After competitions, athletes frequently engage in rigorous physical conditioning and specific skill training. Regular training sessions for athletes are typical of a greater intensity. Athletes' physical and emotional health are harmed by long-term, high-intensity exercise, which leads to sports injuries. An injured athlete could not achieve the expected result in the competition due to his inability to undergo high-intensity training [5]. Based on the relevance of neural networks for target recognition, they present a novel dual-level feature fusion neural network (NN) model for sports injury estimate. The proposed model improves effective discrimination by building a dual-fusion structure with a 1:1 convolution and linkage to overcome the problem of feature loss. According to the trials, the suggested model exhibited a classification accuracy of 97.0 percent, a sensitivity of 95.70 percent, and a specificity of 97.54 percent [6]. The proposed model outperforms all of the other models that were taken into account for this research. Using a neural network to build a sports match prediction model and anticipate the outcomes can provide a theoretical framework for practice, prediction, and analysis. Various data mining techniques examine these medical records [7].

The Recurrent Neural Network (RNN) is a reliable technique that has been extensively used in machine learning and bioinformatics. The research on injury classification and level identification of the spinal cord is done using CT scan images that have been segmented using adaptive thresholding methods. The disc is then localized using the sparse fuzzy C-means clustering technique once the segments have been created. The next stage recovers the connectivity characteristics, statistical features, image-level features, grid-level features, Histogram of Oriented Gradients (HOG), and Linear Gradient Pattern (LGP) from the localized disc. After that, a Deep Convolution Neural Network based on the Crow Search Rider Optimization approach is used to identify damage (CS-ROA-based DCNN). Once the occurrence of the injury has been identified, the proposed Deep Recurrent Neural Network (Deep RNN) is used to categorize the injury severity, and the procedure is terminated if there is no damage [8]. The damage level is determined by the injury detection classifier, which might be normal, wedge, concavity, or crush [9].

A vast body of research has paved the way to improvements in legislation and policies to lessen the incidence and impact of concussions during the previous decade. Young individuals who engage in high-risk activities like sports, on the other hand, commonly underreport concussions, while others may exaggerate reports for a variety of reasons. Such laws and legislation must function within a supportive social context to be effective. Therefore, understanding the culture around concussion is crucial to decreasing concussion and its consequences [10]. An automated deep neural network approach is performed to analyze public opinion on concussion concerns in sports. As a result, one of the project's objectives is to develop a viable method for analyzing and measuring the public's current attitude toward sports-related concussions by examining an extensive collection of public views and perspectives on the subject. Various stakeholders may use this method's data to measure public opinion and attitudes around concussions. This knowledge might aid in the successful implementation of concussion-reduction strategies. Continuous monitoring of lab data and high-dimensional vital signs is needed to detect and diagnose acute patient issues, whereas continuous monitoring has been a difficult task in the past [11]. Recurrent Neural Networks (RNNs), a type of deep learning model, have lately demonstrated their capacity to predict such events. In actuality, however, patient data are continually being added, and RNN does not have a feasible adaptation mechanism to integrate this new data and enhance accuracy. The suggested method relies on the previous projected output and its associated label to update the RNN's hidden state. The additional information offered by the predicted past production and label improved the model's performance. This model's regularization technique considers the model's label prediction errors and its input data estimate errors [12]. The regularization approach reduces the variance of the model and the time gap of the self-correcting mechanism. The proposed model is suitable for both classification and regression. The proposed models were evaluated using real-world, large-scale ICU datasets. This method includes historical timestamp prediction mistakes into the current timestamp forecast, allowing the model to "learn" from previous forecasts. A regularization technique took into account the model's estimating errors on the input data and its label prediction errors [13]. As a result, an intelligent system such as Recurrent Neural Networks (RNNs) is required to efficiently analyze, predict, and recognize sports injuries.

3. Motivation of the Work

Physical education is recognized as an essential component of health education as well as an ability to reflect of a school system. A sports leader is in charge of students' physical education and/or health training but also treatment.

Through student physical education classes at higher education institutions, this study suggested a technique for training but also treatment evaluation. The framework proposed by humans can improve the overall generalizability of the conventional BP network along with training time by integrating a neural network with the BP algorithm.

4. Proposed Model

Sport injuries can sometimes become more dangerous that it can cause disability to the sportsperson. This ends the career of the sportsperson and also causes big loss to the team who invested huge time and money on the particular player. Thus, a frequent monitor of their health is done periodically and the data are being recorded in a database. Each player/athlete has a separate kind of body health. In this scenario, one person's health data cannot be taken as a reference in treating other person beyond certain extent. Thus, the data of a player have to be compared and analyzed with his own dataset that was recorded previously. The data are collected and analyzed with the help of artificial neural networks. In this case, a Recurrent Neural Network (RNN) model has been used to analyze the effect of physical training and treatment in relation with sports injuries is depicted in the Figure 1.

The main phases involved in this method are as follows.

(1) Physical Training: The physical training plays an important part in the career of a sportsperson. Without training session, the fitness level to play the game is lost even by star sportsmen. But over indulgence in training can cause physical stress to the players, leading to sports injuries like muscle tear. Therefore, training has to be done in a manner that the occurrence of sports injury has also to be avoided by predicting it beforehand and taking leave for the session or sport until the completion of the healing process.

(2) Recurrent Neural Network: In order to predict the occurrence of a sports injury, an artificial neural network is assigned to accomplish the task. The ANN used here is Recurrent Neural Network (RNN), which analyzes the neural nodes with respect to the temporal sequence. The temporal sequence is a data collection method of collecting temporal data according to the sequence of happening in a given set of timeframe. These data are analyzed for predicting the sports injury.

(3) Analyze and Predict Sports Injury: There is a strong need in predicting the occurrence of sports injury to avoid the loses that occur in one or more ways. In Recurrent Neural Network (RNN), the outputs of all the neurons are connected to the inputs of the neurons. The Recurrent Neural Network is used as a computational device system in analyzing the data to produce computational data output, which can be used as a predictive measure to avoid sports injuries.

(4) Precaution Measures: With the help of the RNN output, the player is advised to take a break until complete healing process is done. This is also checked with the help of the medical support team and AI tools. Regarding AI tools used in the process, there are a number of AI sensors and biosensors that are already in use in the medical industry. Those biosensors are used to collect data in some cases. Thus, an intelligent RNN plays an important role in predicting the player's health to avoid sports injuries.

5. Proposed Work

The processing elements of such woman are x, which appear to be the analysis value of education physical quality education; the continental surface modules are x, and so are the extracted features i which appear to become the analysis significance of training physical quality education. Since the input image device transmits data directly to a frame node's center, the output of a input layer network equals a input; the information or data of a mid-layer access point seem to be the major contributors of an output node; and also the activation function alone has one network device, which also obtains the information of an intermediate layer access point but also produces an output its training but also treatment requirement evaluation effects.

Input data $m_i, i = \{1, 2, \ldots, x\}$, where x reflects training and treatment quality evaluation.

The equation (1) is represented by the input to the H middle of the specimen node.

$$H_j = \sum_{j=1}^{x} \varphi_{ij} m_i. \tag{1}$$

The result is as follows from the equation (2).

$$R_j = \sum_{j=1}^{x} \frac{1}{\left\{ 1 + \left[\left(\sum_{i=1}^{x} \varphi_i R_j \right)^{-1} - 1 \right]^2 \right\}'} = \sum_{j=1}^{x} \frac{1}{\left[\left(H_j^{-1} - 1 \right)^2 \right]}. \tag{2}$$

In which φ_j signifies the resilience from an input datatype ground station H_j^{-1} to a frame datatype center j as well as R_j signifies the data's factor, the ith training as well as treatment quality evaluation index.

Node of input signal: There are only S nodes with in destination node (as shown in equation (3)), and also the information has been the high speed of the structure node's center point:

$$S = \sum_{j=1}^{x} \frac{1}{\left\{ 1 + \left[\left(\sum_{i=1}^{x} \varphi_i R_j \right)^{-1} - 1 \right]^2 \right\}^{2'}}. \tag{3}$$

The learning optimization technique is defined as even the mean score of such mean square sum $\sum_{i=1}^{x} \varphi_j R_j$ of such an error between such overall performances but also the measuring device value of G measurement techniques and the calculation is displayed in equation (4).

Recurrent Neural Model to analyze the effect of physical training and
treatment in relation to sports injuries
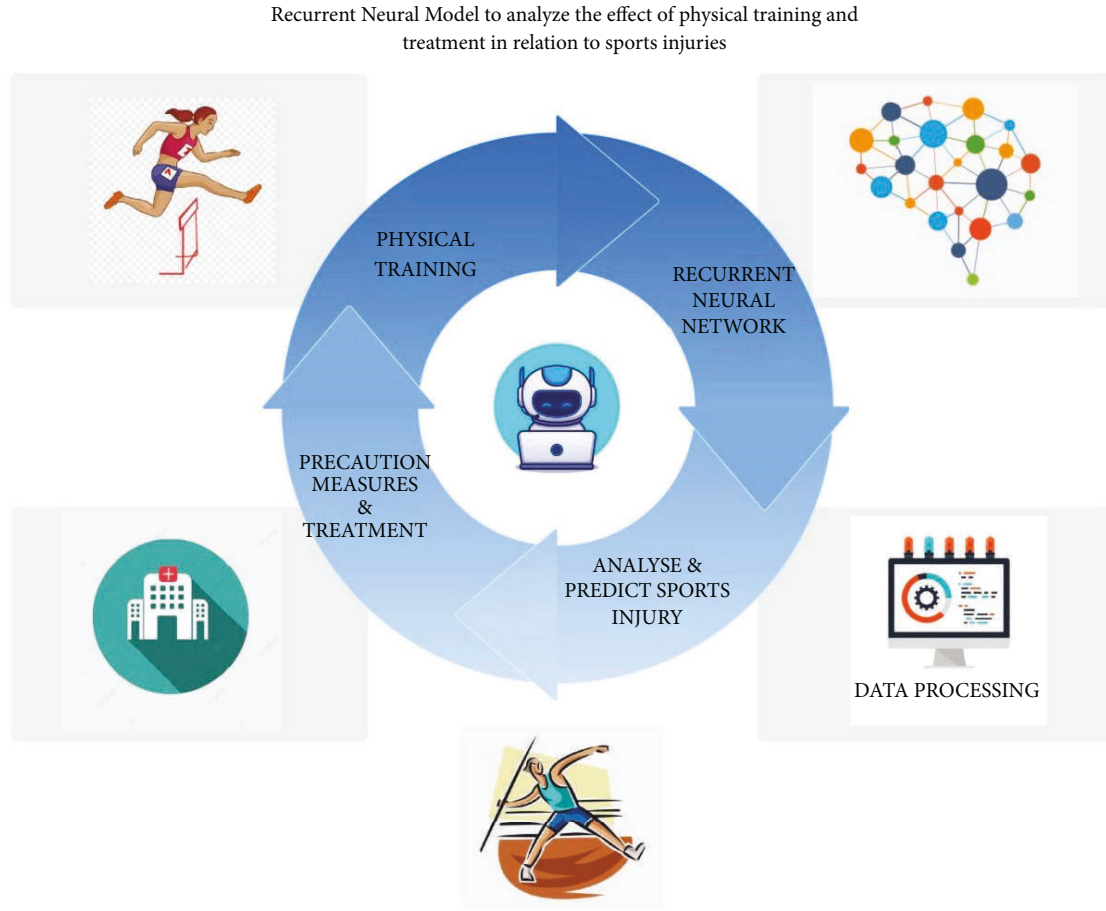
FIGURE 1: Proposed recurrent model for prediction and analysis of sports injuries.

$$G = \left(\frac{1}{M}\right) \sum_{m=1}^{m} \left[\underline{s} - s\right]^2 = \left(\frac{1}{M}\right) \sum_{m=1}^{m} G_j. \qquad (4)$$

Despite the BP neural network evaluation process's modifying δG framework, the goal of network training and treatment is to minimize τ by modifying the network's authentication and authorization. The training algorithm technique δc_{ij} is being used to modify its delinking as in the following equation (5).

$$G_{ij} = \sum \left\{ \varphi_{ij} = -\tau\left(\frac{\delta G}{\delta c_{ij}}\right), \varphi_j = -\tau\left(\frac{\delta G}{\delta \varphi_{ij}}\right). \qquad (5)$$

Furthermore, for such a rate of learning, the $d_i \varphi_j R_j^2$ quantity of communication optimizing parameters here between the information data network and the medium layer wireless device is then shown in the equation (6).

$$\varphi_{ij} = d_i \varphi_j R_j^2 \left[1 - \sum_{i=1}^{x} \varphi_{ij} d_j \right] \omega_j. \qquad (6)$$

The number of connection optimizing parameters is as in the following equation (7).

$$\varphi_i = \sum_{j=1}^{r} S^2 R_j \left[1 - \sum_{j=1}^{r} \varphi_j R_j \right] \left[\underline{s} - s\right]^2. \qquad (7)$$

So, using this framework, its neural network's data transmission weight can also be described to use the $\sum_{j=1}^{r} \varphi_j R_j$ optimization technique of such a specific brain network and the imperfection between these overall performances, and also various sample values can be reduced. The framework will be in a higher optimization procedure, as shown in the equation (8), by using an optimized neural network.

$$(G) = \sum_{j=1}^{r} \int (\varphi_1, \ldots \ldots \varphi_n). \qquad (8)$$

For which (G) seems to be the overall lack of accuracy of network training and $\varphi_1, \ldots \ldots, \varphi_x$ are the constant weights now since strong and secure numbering includes the weights of the network of input access to data points along with center layer access points and also the prediction models of concentrators and output node devices, and n represents the number of network weights. Among most of them, $\underline{s}$ and s are φ_1 variables which it represent its upper and lower bounds of change.

The optimization algorithm is a low capacity problem in the process of analytical algorithms. Because an individual's areas of interest are roughly equivalent to $\int(\varphi_1, \ldots, \varphi_x)$ ability, its representation of an optimization technique has a significant effect on genetic algorithms. Because of its close route planning friendship between both the $U - G$ optimization procedure and the optimization techniques, the $G < U$ strength training computational process described below equation (9) is being used.

$$\int_i = \sum_{i=1} \{U - G \quad G < U, 0 \quad G \geq U. \tag{9}$$

In which e represents the training optimization technique and G represents the sum of all $G \geq U$ in the current generation. The recommendations below are used by the M_r classification of genetic process parameters in order to accept responsibility for such $\int (U - G)/U \in [0, 0.6]$ efficiency of integration and also to eliminate excess integration caused by important evolutionary declassification.

$$M_r = \left\{ 2 \left(\frac{\int_1 U - G \quad G < U}{U} \right) \right\}, \frac{\int_1 (U - G)}{U} \in [0, 0.5], \left[1 - 2 \left(1 - \left(\frac{\int_1 U - G \quad G < U}{U} \right)^2 \right) \right], \frac{\int_1 (U - G)}{U} \in [0.5, 1]. \tag{10}$$

The min and max technique is used for normalization handling $\int_1 (U - G)/U \in [0.5, 1]$ since it is a successful implementation for processing information that can successfully retain its very own original definition while having caused no redundant data. The normalization equation (10) is to use this report for such input data.

The procedures of compacting a large variety of information into the field of view [0, 1] are known as normalization $\int_1 (U - G)/U \in [0.5, 1]$. The solution to the normalization is given in the equation (11) and is as follows.

$$d' = \sum_{i=1}^{x} \frac{d - d_{\min}}{d_{\max} - d_{\min}}. \tag{11}$$

The standardization process σ entails converting the dataset's small and big outlier documentation into a discrete random variable with such an average overall value of 0 and then a confidence level of 1. The following equation (12) depicts the standardization process.

$$d' = \sum_{i=1}^{x} \frac{d - \underline{d}_{\min}}{\sigma}. \tag{12}$$

Each data center is composed of three layers: input nodes, hidden nodes, and also convolution layers $\beta, \gamma, and \, \alpha$, also with weight training of each surface being and also including. It suggests that specific $\varphi_{y'y} q_{y'}^{t-1}$ information is managed to retain in RNN receptors after each cycle of data transmission. It must enter a $\varphi_y (q_{y'}^{t-1})$ next nerve cell as new information and effect the subsequent data output. The equation (13) for such corresponding input nodes, its original input of both the hidden units, and also the outcome variable to it of destination node at time step t is represented in the following: equations (14) and (15).

$$\beta_y^t = \sum_{i=1}^{M} \varphi_{iy} d_i^t + \sum_{y'}^{M} \varphi_{y'y} q_{y'}^{t-1}, \tag{13}$$

$$\alpha_{y'}^{t-1} = \varphi_y \left(q_{y'}^{t-1} \right), \tag{14}$$

$$\gamma_{y'}^{t-1} = \sum_{y=1}^{M} \varphi_{y0} q_y^t. \tag{15}$$

As a result, by modeling differently in time-distance, this framework is enhanced much further, seeking to make this more distinctive than spatial transformation. The distributed probability, which gets to know to extract temporal features from streaming video using learning algorithm, is being used in the case as part of a hierarchy approach. As the introducing module, its own interpolation constricted machine effort in learning the hierarchy organization of the original data structure. The framework has become more difficult since it develops from highest to lowest level. The geographically but also temporally High Duplicate Network is named after a constant growth in interpretations.

Previously, the rapacious hierarchical model was used to train the Enlarged Deep Learning (Deep Belief Network (DBN)) model. Furthermore, each framework's input layers are informed at random, beginning with the lowest layer. The probability recognition of the hidden nodes is then reorganized, and understanding to some other layer is gained. This process was repeated indefinitely during training until all of the layers have been trained, since training the entire network might recover the hidden and exposed possibility representations of every particular part within the video.

6. Results and Discussion

The above graph having to learn optimization method is defined as the average score of these mean square $\sum_{i=1}^{x} \varphi_j R_j$ of such a mistake between such overall performance and the measuring handset value of G measurement methods. As in equation (4), Figure 2 depicts the physical education training and treatment performance of fourth- and eighth-grade students. The effectiveness of training and treatment is evaluated using a Recurrent Neural Network Model, a Support Vector Machine, and a Fuzzy Set Model based on hesitation. The computation is carried out by combining the
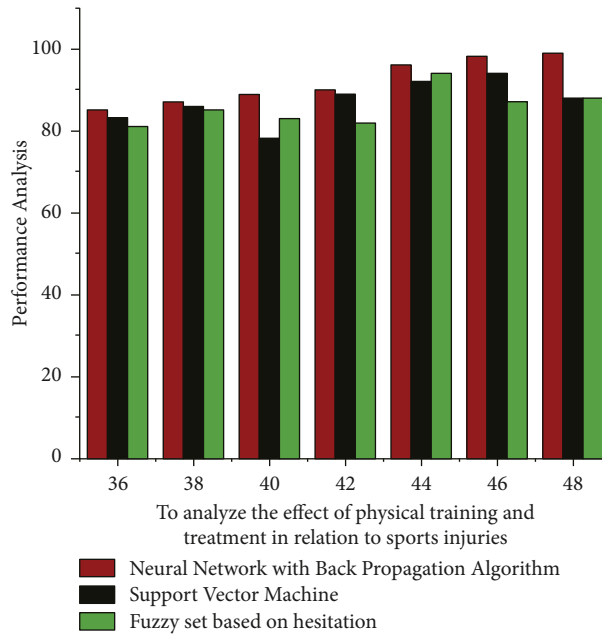
FIGURE 2: The computation accuracy of measurement weight results is compared.

TABLE 1: Analyze the effect of physical training and treatment on sports injuries using a result analysis. Precision in computing.

Effect of physical training and treatment in relation to sports	Recurrent Neural Network model (%)	Support Vector Machine (%)	Fuzzy set based on hesitation (%)
36	85	83	81
38	87	86	85
40	89	78	83
42	90	89	82
44	96	92	94
46	98	94	87
48	99	88	88

performance of the fourth- and eighth-grade students. The proposed algorithm, Recurrent Neural Network Model, has a lower percentile at the early stage of new technology. However, at a later stage, the algorithm was able to produce results that were comparable to the Support Vector Machine. The fuzzy model yields results that are indistinguishable from the Recurrent Neural Network Model and Support Vector Machine (SVM) algorithms. Such a comparative graph demonstrates how repeated training, treatment, and learning can improve the performance of students and teachers when using the Recurrent Neural Network Model. Table 1 shows that its Recurrent Neural Network Model provided small training accuracy than other algorithms for an average evaluation weight of 38. Furthermore, the accuracy rate of the Recurrent Neural Network Model is steadily improving compared to the other two algorithms with fluctuating results. Later on, the proposed algorithm achieved an accuracy of 99%, which is a minimum increase of nine percentage points over the Support Vector Machine and then a 18% increase over the Fuzzy set model.

The δG framework is being modified as part of the BP neural network evaluation process. The goal of capacity training, as well as treatment, seems to be to minimize τ by modifying Internet backbone authentication process. To adjust its delinking, the developer these skills technique δc_{ij} is used resemblance of performance measures with those of algorithms represent in Figure 3 also demonstrates that such built model performance when compared conventional algorithms, including such neural network models, through seeking to resolve activity recognition problems. The 98.55% training and treatment accuracy rate also confirms its high specificity. The convolutional neural network could capture the qualities of training (93.26%), training and treatment (95.47%), accuracy and loss of training (88.74%) analysis, optimize the required data, and improve the consistency of exercise physiology recognition simply by attempting to adjust the set of parameters. The neural network accuracy training and testing graph is capacity forecast in physical education training and treatment based on BP neural network.
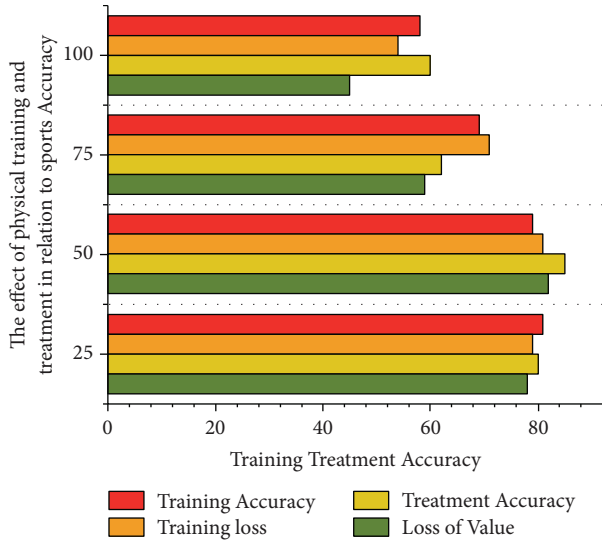
FIGURE 3: The effect of physical training and treatment on neural network accuracy but also loss in sports using the BP algorithm.

Moreover, with such a learning algorithm, the $d_i \varphi_j R_j^2$ quantity of communication optimization parameters between both the information data connection and the medium surface wireless device is shown in equation (6) elsewhere here represented in Figure 4; it also shows that trained human activity recognition framework appears to have an evaluation of overall accuracy of much more than 99.7%, precision of much more than 92.46 percent, recall frequency of many more than 89.72%, for the different topics; until training, the activity recognition framework for several Physical Education.

The isolation of visual medium, which includes higher rates of physical and also postures judgment in physical training and treatment, has been shown to continue providing educators with unbiased information about student movements, and thus, this recognition of educational body actions may provide legitimate input to improve training performance. Information from intelligent smart wearable sensors has been used to accurately determine human behavior, but the classification of constructions is a controversial subject. The experimental results (refer to Table 2) show that such a convolutional network-based activity recognition system could indeed accurately recognize human behavior with an accuracy of more than 98%. As a result of the Analysis for Multiple Physical Education assessment set in diverse physical education for precision (89%), recall (87%), and accuracy (99%), as well as the Human Sports Activity of precision (90%), recall (85%), and accuracy (98%).

The suggestions below are being used by the M r categorization of genetic process variables to take accountability for such $\int_1 (U - G)/U \in [0, 0.6]$ integration efficiency and to remove unnecessary integration caused by significant evolutionary release of information. The evaluation process of physical education courses follows the training, treatment, and learning processes and the analysis is shown in Figure 5.
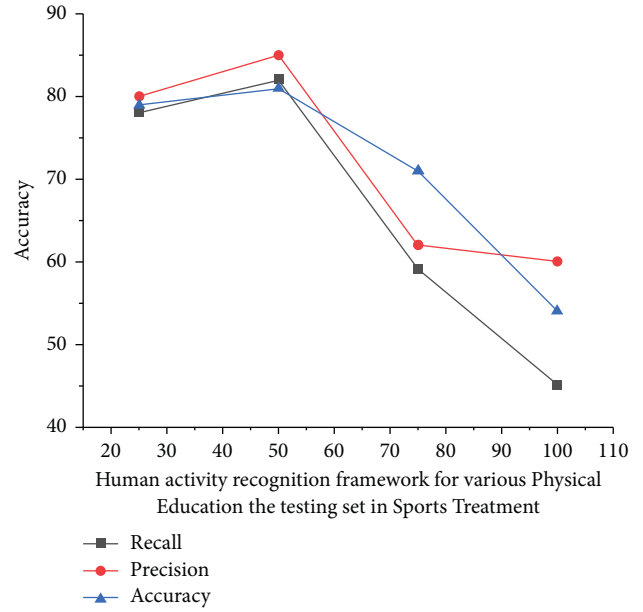
FIGURE 4: Performance evaluation of human activity recognition.

TABLE 2: Analysis for the result indifferent physical education the training set in sport treatment.

	Precision (%)	Recall (%)	Accuracy (%)
Various physical education	0.88	0.89	0.99
Human sports activity	0.91	0.86	0.97
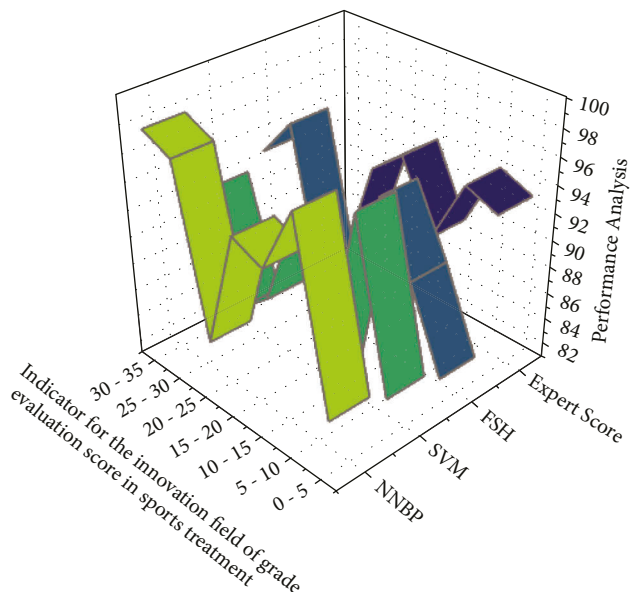
FIGURE 5: Analysis of evaluation score comparison results in sport training and treatment in physical education.

TABLE 3: Comparison result analysis for different algorithm in the effect of physical training and treatment in relation to sports injuries.

Effect of physical training and treatment	Recurrent Neural Network model (%)	Support Vector Machine (%)	Fuzzy set based on hesitation (%)	Expert score
0–5	85	84	83	93
5–10	98	96	89	94
10–15	93	83	94	89
15–20	94	93	86	94
20–25	85	86	85	92
25–30	97	84	96	84
30–35	98	92	93	82

When compared to expert evaluation, the proposed automatic evaluation via the Recurrent Neural Network model outperformed many assessments. The graph's peaks represent the evaluation value for the specific assessment for students in both grades. Table 3 shows the numerical representation. According to this table, the proposed Recurrent Neural Network model demonstrated fluctuating results during the learning process; however, it demonstrated consistent and increased performance at a later stage. Towards the final considered evaluation weight range of 30–35, the proposed system outperformed the existing Support Vector Machine, fuzzy set predicated on doubt, but also expert score by 6%, 5%, and 16%, respectively.

Figure 6 depicts the analysis of results performed on students of various grades again for physical education training, treatment, but also learning process. Students are encouraged to practice on a regular basis in order to be eligible for competitions in any sport. The graph above shows that the analysis is performed on people from four different categories, including normal people, majors, juniors, and seniors, using sports items, competition, and practice as parameters. All of the necessary equipment for practicing and participating in any specific sport is included in a sports item. In this study, it is assumed that the player has all of the necessary resources. Among some of the four categories of people, the average person with no prior experience was only able to achieve 16% in the competitions despite having access to all of the necessary resources. Individuals in the major, junior, but also senior categories, on the other hand, perform better by 23%, 35%, and 43%, respectively. Every data center is made up of three layers: input endpoints, hidden nodes, and convolution layers $\beta, \gamma, and \alpha$ with weight training of every surface included. It implies that specific $\varphi_{y'y} q_{y'}^{t-1}$ information is managed to maintain in RNN receptor sites after each data transmission cycle. It must enter the next neuron as new information and affect the subsequent data output.

The data in Figure 7 above demonstrate that training is required to participate in any sport, and learning them from a coach is also very important. The number of indicator or exercises increases with changes in this graph, as does the computing time or the players' duration of achievement. Initially, each person will devote the maximum amount of time to simple and short exercises. The duration of computation will be reduced as the players perform workouts and practices on a regular basis. It must enter the next neurone as new information as well as affect the subsequent
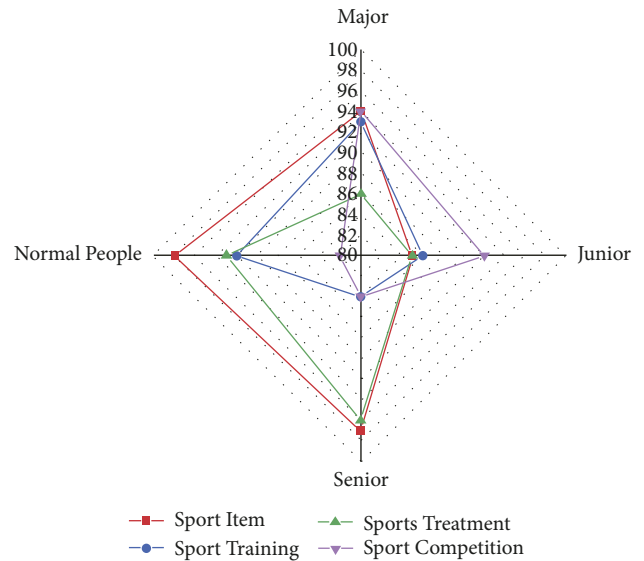
FIGURE 6: Analysis for result in the physical training and treatment.
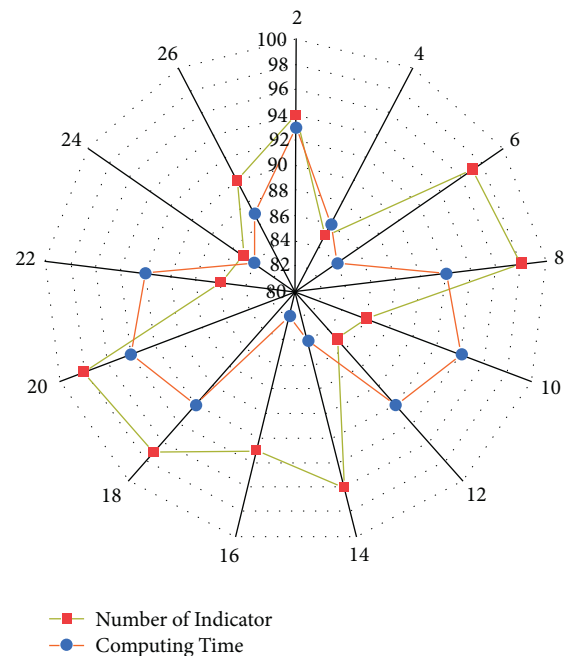
FIGURE 7: The thorough outcomes of sports training and treatment are critical.

data output. Figure 7 depicts the equations (13), for such corresponding input nodes, its original input of both hidden units, and the outcome variable to it of the destination node at time step t.

7. Conclusion

Injuries are most common among the sports persons during practice sessions and event participation. It is mandatory for them to recover from the injuries at the earliest and start a new beginning. In the technological development, physical training and treatment for the sports injuries can be monitored with the implementation of the intelligent system. This advancement in the technology aids the sports person and the coach to make regular updations on the status of improvement on the injuries concerning to the practice and treatment. In this research, Recurrent Neural Model is designed to perform frequent monitoring of the processes involved in sports injuries and predict the possibilities of injuries. The proposed model outperformed the existing Support Vector Machine, fuzzy set predicated on doubt, but also expert score by 6%, 5%, and 16%, respectively.

Data Availability

The data used to support the findings of the study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] L. Meng and E. Qiao, "Analysis and design of dual-feature fusion neural network for sports injury estimation model," *Neural Computing & Applications*, 2021.

[2] P. Chantamit-o-pas and M. Goyal, "Long short-term memory recurrent neural network for stroke prediction," in *Proceedings of the International Conference on Machine Learning and Data Mining in Pattern Recognition*, pp. 312–323, Springer, Cham, 2018, July.

[3] K. Tirdad, A. Dela Cruz, A. Sadeghian, and M. Cusimano, "A deep neural network approach for sentiment analysis of medically related texts: an analysis of tweets related to concussions in sports," *Brain Inf*, vol. 8, no. 1, p. 12, 2021.

[4] H. Du, Z. Pan, K. Y. Ngiam, F. Wang, P. Shum, and M. Feng, "Self-correcting recurrent neural network for acute Kidney injury prediction in critical Care," *Health Data Science*, vol. 2021, Article ID 9808426, 2021.

[5] M. Kallinen and A. Markku, "Aging, physical activity and sports injuries," *Sports Medicine*, vol. 20, no. 1, pp. 41–52, 1995.

[6] X. Wang and C. Jiang, "Computer-aided physical training sports injury risk simulation based on embedded image system," *Microprocessors and Microsystems*, vol. 83, Article ID 103975, 2021.

[7] A. Zadeh, D. Taylor, M. Bertsos, T. Tillman, N. Nosoudi, and S. Bruce, "Predicting sports injuries with wearable technology and data analysis," *Information Systems Frontiers*, vol. 23, no. 4, pp. 1023–1037, 2021.

[8] D. A. Dobrosielski, L. Sweeney, and P. J. Lisman, "The association between poor Sleep and the incidence of sport and physical training-related injuries in Adult athletic Populations: a Systematic Review," *Sports Medicine*, vol. 51, no. 4, pp. 777–793, 2021.

[9] S. Kozin, Z. Kozina, V. Korobeinik et al., "Neuro-muscular training for injury prevention of students-rock climbers studying in the specialty "Physical Education and Sports": a randomized study," *Journal of Physical Education & Sport*, vol. 21, 2021.

[10] S. Xianguo and W. Cong, "Research on the Application of artificial intelligence technology in physical training," in *Proceedings of the 2021 2nd International Conference on Big Data and Informatization Education*, pp. 261–264, ICBDIE, November 2021.

[11] D. Li, C. Yi, and Y. Gu, "Research on College physical education and sports training based on Virtual Reality technology," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–8, 2021.

[12] N. Coleman, "Pediatric athlete development and Appropriate sports training," in *Common Pediatric Knee Injuries*, N. Coleman, Ed., Springer, Cham, 2021.

[13] O. V. Balberova, "Candidate genes and single-nucleotide gene variants associated with muscle and tendon injuries in cyclic sports athletes," *Personalized Psychiatry and Neurology*, vol. 1, no. 1, pp. 64–72, 2021.

Research Article

Alterations of Renal Function in Patients with Diabetic Kidney Disease: A BOLD and DTI Study

Xiaobao Wei [ID],[1] Runyue Hu,[2] Xiaoli Zhou,[2] Lihua Ni,[1] Dongqing Zha,[1] Huiling Feng,[1] Haibo Xu [ID],[2] and Xiaoyan Wu [ID][1]

[1]*Department of Nephrology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China*
[2]*Department of Radiology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China*

Correspondence should be addressed to Haibo Xu; xuhaibo1120@hotmail.com and Xiaoyan Wu; wuxiaoyan2k6@163.com

Objectives. Our study aims to determine the patterns of renal oxygenation changes and microstructural changes by BOLD and DTI with deteriorating kidney function in patients with diabetic kidney disease (DKD). *Methods*. Seventy-two patients with type 2 diabetes mellitus (DM) and twenty healthy controls (HCs) underwent laboratory examinations, and renal BOLD and DTI images were obtained on a 3T-MRI machine. $R2^*$, fractional anisotropy (FA), and average diffusion coefficient (ADC) values were evaluated. DM patients were divided into three subgroups (Group-DI/DII/DIII, based on urinary albumin-creatinine ratio (UACR)) and a nondiabetic kidney disease group (Group-NDKD). D-value and MCR of $R2^*$ and FA were proposed to evaluate the differentiation between medulla and cortex of the individual kidney among HCs and three subgroups for reducing individual differences. Comparisons were made between NDKD and kidney function-matched DKD patients. Correlations between MRI parameters and renal clinical indices were analyzed. *Results*. Compared with Group-HC/DI, medullary $R2^*$ and FA values were significantly different in Group-DII/III. The D-value of $R2^*$ and FA in Group-III were significantly smaller than that in Group-HC. However, only MCR of $R2^*$ in Group-III was significantly smaller than that in HCs. Medullary $R2^*$ and FA were negatively associated with serum creatinine (SCr) and cystatin C (Cys C) and positively associated with eGFR. *Conclusions*. With renal function declining, BOLD and DTI could capture alterations including the first rising and then falling medullary $R2^*$, continuously declining medullary FA, and apparent cortex-medullary differentiation in DKD patients. The MRI parameters showed renal changes accompanied by varying degrees of albuminuria, sharing common involvement in DKD and NDKD patients, but it was hard to distinguish between them. BOLD seemed more sensitive than DTI in identifying renal cortex-medullary differentiation.

1. Introduction

Diabetes usually results in complications, affecting the nerves, eyes, kidneys, and cardiovascular system, which leads to decreased quality of life and a greater risk of death than similarly aged healthy people [1]. DKD has been reported to have an incidence of 20%∼40% in diabetic patients in European and American countries [2, 3]. In 2011, the prevalence of DKD in hospitalized patients in China began to exceed that of glomerular nephritis-related chronic kidney disease and the gap has been increasing [4, 5].

The American Diabetes Association (ADA) and Kidney Disease: Improving Global Outcomes (KDIGO) group suggest that the detection of the urinary albumin/creatinine ratio (UACR) and eGFR should be appropriately frequent to better diagnose and treat DKD [6, 7]. The clinical symptoms of early DKD are not obvious. Patients may have dry mouth, anorexia, stomach upset, fatigue, weakness, mild eyelid edema, or even numbness, but usually do not have symptoms such as moderate edema, nausea, and vomiting. When urinary microalbumin (uAlb) or SCr is used as a basic means to predict and screen DKD, once it is abnormal, kidney function already declines significantly [8, 9]. Biomarkers, including indicators of glomerular damage and tubular damage, inflammatory factors, proteomics, and even microRNAs, require in-depth clinical validation [10].

Moreover, the renal lesions of DKD cannot be visually displayed by routine ultrasound, CT, and MRI examination.

It is essential to assess kidney damage in DKD patients via a noninvasive and intuitive method. Sufficient evidence indicates that chronic hypoxia and interstitial fibrosis in the kidneys are involved in the occurrence and development of DKD [11–14]. The early pathological manifestations are increased glomerular volume, increased mesangial matrix, and thickened basement membrane, and the late stage is mostly sclerotic changes. Blood oxygenation level-dependent magnetic resonance imaging (BOLD-MRI) is the only noninvasive technology for evaluating the renal oxygen content in renal living tissue instead of invasively inserting microelectrodes into the kidneys directly [15]. In BOLD, deoxygenated hemoglobin is paramagnetic, and its content will affect the uniformity of the magnetic field in the surrounding local tissues, which is transformed into the $T2^*$ signal. Since $R2^* = 1/T2^*$, anything that could reduce the concentration of deoxygenated hemoglobin decreases the $R2^*$ value. An increase in the $R2^*$ value corresponds to a decrease in oxygenated hemoglobin, a decrease in partial pressure of oxygen, and tissue hypoxia. The hyperperfusion of the renal medulla and its oxygen partial pressure gradient allow BOLD-MRI to detect slight fluctuations in deoxy-hemoglobin in the renal medulla, making it easier to evaluate renal oxygenation. Diffusion tensor imaging (DTI) is the only noninvasive technology for evaluating water molecular diffusion in the renal living tissue, which can detect the altered structures in the medulla by abnormal diffusion of water molecules in patients with early DKD [16, 17]. Diffusion refers to the random and irregular movement of molecules, which is an important physiological activity of the human body. Renal DTI mainly describes the range of diffusion movement of water molecules in unit time by ADC and describes the proportion of anisotropic components of water molecules in the whole diffusion tensor by FA to reflect the renal microstructure. The purpose of our research was to explore the relationships between BOLD and DTI alterations and the renal function decline in DKD patients and to explore whether there is a difference between the renal damage exhibited by NDKD and DKD in DM.

2. Materials and Methods

2.1. Participants. This study was approved by the Ethics Committee. All participants signed an informed consent form before participating in the experiment. We recruited 80 patients from the Department of Nephrology and Endocrinology in our hospital and 20 healthy volunteers matched the sex and age criteria from December 2020 to June 2022. The patient's inclusion criteria were as follows: (1) age 18 to 80 years; (2) type II DKD, in accordance with the 2020 American Diabetes Association diagnosis, eGFR >30 ml/min/1.73 m^2; (3) type II simple DM, no renal complications or biochemical abnormalities; (4) nondiabetic kidney disease (NDKD) patients (type II diabetes with CKD, e.g., hypertensive nephropathy, nephrotic syndrome, and tubulointerstitial disease), eGFR >30 ml/min/1.73 m^2; (5) no serious complications of diabetes, such as diabetic ketosis,

hyperglycemic hyperosmolarity syndrome, and hypoglycemia; no diuretic use within one month and no renal replacement therapy; and (6) no MRI contraindications. Eight cases were completely excluded because of multiple or large kidney cysts (the presence of a cyst larger than 4 cm in diameter in one kidney or the presence of multiple cysts with the sum of their diameters larger than 4 cm, $n = 6$) and eGFR <30 ml/min/1.73 m^2 (eGFR too low, no control, $n = 2$). Finally, 62 T2DM and DKD patients, and 10 NDKD patients were included in this study. 62 patients were divided into three subgroup according to UACR: DI group (UACR ≤ 30 mg/g); DII group (30 < UACR ≤ 300 mg/g); DIII group (UACR> 300 mg/g), and NDKD group (30 ml/min/1.73 m^2 < eGFR < 90 ml/min/1.73 m^2).

Healthy volunteers were included accordant with the following specific criteria: no history of urinary system disease, no diabetes, and no use of any drugs that damage the kidneys; blood pressure control within 140/90 mmHg; no abnormalities in routine blood tests and liver and kidney function tests; good cooperation with MRI examinations; and no MRI contraindications.

2.2. Study Design. We used a 3.0-T machine (Signa 750w, GE Healthcare, USA) with a 16-channel body coil. Before the scans, all the subjects were abstained from consuming food and water for 4 hours and underwent breathing training. Routine $T2$ and $T1$ sequences were scanned after localization to exclude significant renal lesions.

2.3. Scanning Parameters. Coronal renal BOLD scanning: we used a multiple fast gradient recalled echo (MFGRE) sequence with 8 echoes with the following parameters: TR = 100 ms, TE = 3.3–26.2 ms, FOV = 32 cm × 32 cm, NEX = 1, slice thickness = 5 mm, spacing = 0.4 ms, matrix size = 256 × 256, bandwidth = 62.50 kHz per pixel, and flip angle = 20°. Images were acquired during a 14 second breath hold with an acquisition time of approximately 27 s, and the central line was placed at the renal hilum, with five layers in front and behind.

Coronal renal DTI scanning: images with respiratory triggering were acquired with a diffusion gradient oriented in 16 directions, b-values of (0 s/mm^2, 600 s/mm^2), using a spin-echo (SE) echo-planar imaging (EPI) sequence with the following parameters: TR = 2000 ms, TE = 66–462 ms, FOV = 36 cm × 32 cm, NEX = 1, slice thickness = 5.0 mm, spacing = 0 ms, matrix = 128 × 128, bandwidth = 250 kHz per pixel, acquisition time = 36–47 seconds, and a range covering both kidneys.

2.4. Image Analysis. We measured the renal cortical $R2^*$ ($CR2^*$), medullary $R2^*$ ($MR2^*$), renal cortical FA (C.FA), renal medullary FA (M.FA), cortical average diffusion coefficient (C.ADC), and medullary average diffusion coefficient (M.ADC). All data were processed by two well-trained doctors independently with three years and six years of experience in imaging diagnosis, who were blinded to the population group before the measurement. We selected

layers close to the renal hilum and manually drew regions of interest (ROIs) on BOLD and DTI images at the upper, middle, and lower poles of the cortex and the medulla of each kidney with ellipses of approximately 40–60 mm^2. We tried to avoid renal blood vessels, edges, collecting systems, and cysts. Examples of pseudocolor images of the $R2^*$ map, FA map, and ADC map are displayed in Figure 1.

2.5. Laboratory Inspection. We collected biochemical data before or after the MRI examinations while excluding the effects of exercise, infection, fever, congestive heart failure, menstruation, significant hyperglycemia, and significant hypertension within 24 hours [18]. We calculated the eGFR based on the Chinese Modification of Diet in Renal Disease (MDRD). The improved formula for eGFR is as follows: eGFR (mL/min/1.73 m^2) = 175 × (serum creatinine)$^{-1.154}$ × (age)$^{-0.203}$ × (0.742 female) [19, 20].

2.6. Statistical Analysis. Statistical analysis was performed using IBM SPSS Statistics 26.0, and the statistical significance was defined as $p < 0.05$. The parameters are expressed as the mean ± standard deviation (SD). One-way ANOVAs with the Bonferroni correction were used to test differences in $R2^*$, ADC, FA values among groups as well as D values (Δ, the measured medulla value-the corresponding cortical value) and MC ratios (MCR, the measured medulla value divided by the corresponding cortical value) of $R2^*$ and FA across groups. A Spearman correlation was also computed between the MRI parameters and clinical indices of renal function.

3. Result

3.1. General Clinical Characteristics of All Participants. The general clinical information and laboratory results are shown in (Table 1), including sex, age, hemoglobin (HGB), body mass index (BMI), eGFR, blood urea nitrogen (BUN), hemoglobin A1c (HbA1c), uAlb, UACR, and Cys C. There was no statistically significant difference in sex, age, or BMI among the HCs and the three diabetic subgroups. HbA1c in the three diabetes subgroups showed no difference. The mean eGFR of the DI group (115.25 ± 16.25 mL/min/1.73 m^2) was slightly higher than that of the HC group (109.49 ± 9.70 mL/min/1.73 m^2) ($p > 0.05$). HGB, SCr, BUN, uAlb, and Cys C were all significantly different among the three diabetes subgroups. The differences were mainly between the DI group and the DII/DIII group based on post-hoc multiple comparisons ($p < 0.05$).

Table 2 shows that there was no significant difference in sex, age, and clinical data between the DKD group (30 ml/min/1.73 m^2 < eGFR < 90 ml/min/1.73 m^2) and NDKD group.

Since renal puncture was an invasive method, most patients refused it when they were hospitalized. Among all the recruited subjects, 21 were definitely diagnosed with diabetes kidney disease and 10 more were diagnosed with NDKD by renal pathology. The specific pathological results and staging are shown in Table 3.

3.2. Comparison of $R2^*$, FA, and ADC of the Cortex and Medulla among the HC Group and the Three Diabetes Subgroups. Intraclass correlation coefficient (ICC) at a 95% confidence interval on values was analyzed, such as $CR2^*$ (ICC = 0.908, 95% CI 0.841–0.946), $MR2^*$ (ICC = 0.910, 95% CI 0.844–0.948), C.FA (ICC = 0.886, 95% CI 0.781–0.937), M. FA (ICC = 0.927, 95% CI 0.874–0.958), C.ADC (ICC = 0.929, 95% CI 0.861–0.961), and M. ADC (ICC = 0.898, 95% CI 0.828–0.940). All ICCs were all greater than 0.75. Data processed by the two doctors showed good consistency ($p < 0.05$).

We performed paired-samples t test between the left and right kidneys (Figure 2). There was no difference in cortical and medullary $R2^*$ between the left and right kidneys ($p > 0.05$). However, FA and ADC values of right kidneys were significantly lower in both cortex ($t = -6.054$, $p < 0.001$; $t = -0.530$, $p < 0.001$) and medulla ($t = -4.162$, $p < 0.001$; $t = -3.364$, $p < 0.01$) than those of the left kidneys. Considering that the left kidney was most likely affected by the heart and blood vessels, we decided to individually utilize the DTI data of the right kidney.

$R2^*$, FA, and ADC values of the cortex were not significantly different among the HCs and the three diabetes subgroups ($p > 0.05$) (Table 4).

Moreover, medullary $R2^*$ values in DI group were significantly higher than those in the HC group ($p < 0.001$). Medullary FA values in the HC group were higher than those in DI group ($p = 0.005$). Medullary ADC values were not significantly different among HCs, the DI, DII, and DIII group. We found no difference in $R2^*$, FA, and ADC values between DKD and NDKD ($p > 0.05$) (Figure 3).

To reduce individual differences, we calculated the D values (Δ) and MC ratios (r) between the renal medulla and the corresponding cortex (Figure 4). Relative to the stable differences between the cortex and medulla of the kidneys in the HC group, the changes in the DIII group implied an decreasing functional and structural differentiation of the cortex and medulla as renal function deteriorated.

3.3. Correlative Analysis. Spearman correlation analyzed the relationships between the cortical and medullary $R2^*$, FA values, and the eGFR, BUN, HbA1c, uAlb, and Cys C values among the DM patients (Figure 5). The eGFR had a moderate and positive correlation with the $MR2^*$ value and a positive but weak correlation with the M.FA ($p < 0.001$). In contrast, SCr was moderately and negatively correlated with $MR2^*$ and negatively but weakly correlated with the M.FA ($p < 0.001$). Cys C had a weak and negative correlation with the $MR2^*$ and M.FA ($p < 0.01$, $p < 0.05$). In addition, there was a correlation between SCr and the $CR2^*$ value, but the correlation was weak.

4. Discussion

The present study verified the (patho) physiological differences between the kidney cortex and medulla. Based on the trends in $R2^*$ and FA values, we can get a general idea of the

(a)

(b)

(c)

(d)

(e)
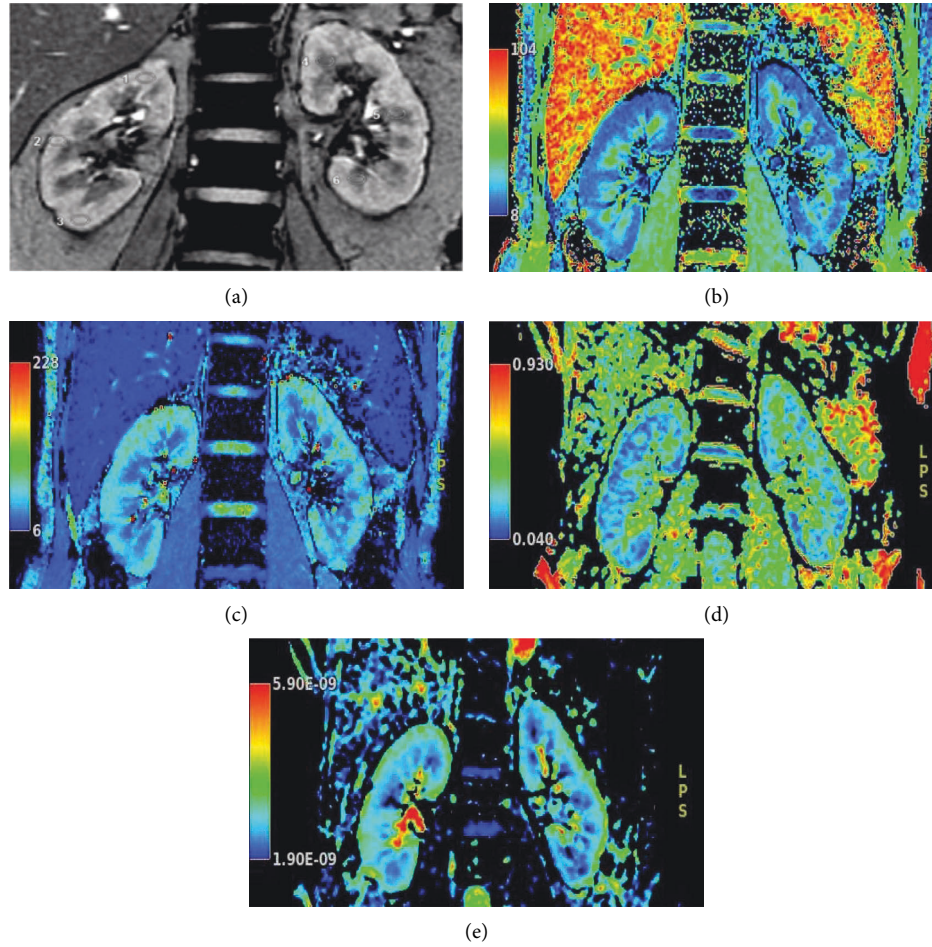
Figure 1: Examples of the positions of ROIs and pseudocolor pictures of BOLD and DTI images. Legends: An example of the measurement of the ROI on the coronal BOLD image (TE = 15.9 ms) in a 49-year-old woman (a). The ROIs were positioned in the cortex (1, 2, and 3) and medulla (4, 5, and 6) in the upper, middle, and lower parts of each kidney. The ROI had an area of 40–60 mm^2. A pseudocolor figure of a participant in the DII group (male, 51 years old) was selected, where $b/c/d/e$ correspond to the $R2^*$, $T2^*$, FA, and ADC images, respectively.

Table 1: Characteristics of the subjects.

Characteristic	HC ($n = 20$)	DI ($n = 22$)	DII ($n = 22$)	DIII ($n = 18$)	F	p
Sex (M/F)	9/11	12/10	14/8	12/6	—	0.510
Age (Y)	54.23 ± 14.69	54.17 ± 14.32	57.75 ± 13.56	58.13 ± 15.68	0.184	0.906
BMI (kg/m^2)	22.49 ± 1.42	24.42 ± 4.60	25.89 ± 2.31	24.65 ± 3.04	2.570	0.068
HGB (g/L)	137.3 ± 8.1	145.9 ± 10.3	133.0 ± 11.3	110.6 ± 19.9	16.804	0.002
eGFR (mL/min/1.73 m^2)	109.49 ± 9.70	115.25 ± 16.25	74.88 ± 7.43	44.44 ± 11.92	75.923	<0.001
Scr (μmol/L)	58.74 ± 5.87	61.83 ± 9.29	84.96 ± 14.00	138.38 ± 45.31	18.114	<0.001
BUN (mmol/L)	5.05 ± 0.92	5.19 ± 0.98	6.92 ± 1.68	7.70 ± 3.14	2.764	0.012
Cys C (mg/L)	0.85 ± 0.18	0.89 ± 0.17	1.12 ± 0.31	1.91 ± 0.68	13.688	0.002
HbA1c (%)	N/A	6.43 ± 0.69	7.03 ± 1.68	6.27 ± 0.72	1.183	0.424
uALB (mg/L)	6.5 (2.8, 12.6)	9.8 (3.2,12.3)	150.0 (28.1, 171.5)	2817.7 (565.3, 3640.3)	11.425	0.038
UACR (mg/g)	12.2 (7.8, 15.3)	13.9 (8.7, 15.8)	58.5 (39.7, 62.5)	2742.0 (567.3, 3813.3)	28.402	<0.001

Legends: The results are expressed as the mean ± SD. If the normal distribution is not satisfied, the results are expressed by quartile method. The chi-square test was used to test the difference in sex in the four groups. ANOVA was used to test difference in the other characteristics among the abovementioned groups. A $p < 0.05$ was considered significant. BMI (body mass index), weight (kg)/height (m^2); HGB, hemoglobin; eGFR, estimated glomerular filtration rate; SCr, serum creatinine; BUN, blood urea nitrogen; uALB, urinary microalbumin; Cys C, cystatin C; HbA1c, hemoglobin A1c; HC, healthy control group; DI, diabetes patients without albuminuria; DII group, diabetic kidney disease patients with microalbuminuria; DIII group, diabetic kidney disease patients with massive albuminuria.

TABLE 2: Characteristics of the DKD group vs. NDKD group.

Characteristic	DKD ($n = 20$)	NDKD ($n = 10$)	T	P
Sex (M/F)	14/6	8/2	—	0.559
Age (y)	57.30 ± 14.05	59.30 ± 10.10	−0.400	0.692
BMI (kg/m^2)	25.43 ± 2.59	25.71 ± 3.03	−0.266	0.765
HGB (g/L)	123.6 ± 18.8	131.8 ± 23.3	−1.028	0.143
eGFR (mL/min/1.73 m^2)	62.71 ± 17.85	59.52 ± 25.70	0.397	0.694
Scr (μmol/L)	106.33 ± 39.88	121.20 ± 47.63	−0.903	0.374
BUN (mmol/L)	7.23 ± 2.33	9.53 ± 7.15	−1.325	0.196
Cys C (mg/L)	1.44 ± 0.62	2.41 ± 2.99	−1.020	0.333
HbA1c (%)	6.75 ± 1.43	6.77 ± 0.99	−0.153	0.880
uALB (mg/L)	1781.2 (1562.3, 1893.5)	1889.0 (1728.4, 2065.3)	−0.171	0.866
UACR (mg/g)	265.6 (243.4, 287.1)	280.4 (236.7, 310.6)	−0.893	0.799

Legends: An independent sample t test was used to compare the other characteristics in two groups. A $p < 0.05$ was considered significant. NDKD group, nondiabetic kidney disease patients; DKD group, diabetic kidney disease patients with matched eGFR.

TABLE 3: Staging of DKD and NDKD patients based on renal biopsy.

Stage of DKD	DKD ($n = 21$)	NDKD ($n = 10$)
Stage I: Glomerular hyperfiltration	$N = 2$	
Stage II: Silent stage	$N = 4$	Diabetic patients with other nephropathy (membranous nephropathy, $N = 2$, hypertensive renal
Stage III: Incipient nephrology	$N = 12$	impairment $N = 3$, glomerulonephritis disease, $N = 3$, tubulointerstitial disease, $N = 2$)
Stage IV: Overt nephrology	$N = 3$	

Legends: The final outcome for staging of patients with DKD took into account both albuminuria and pathological findings. None of the 21 DKD patients had other etiologies of CKD.
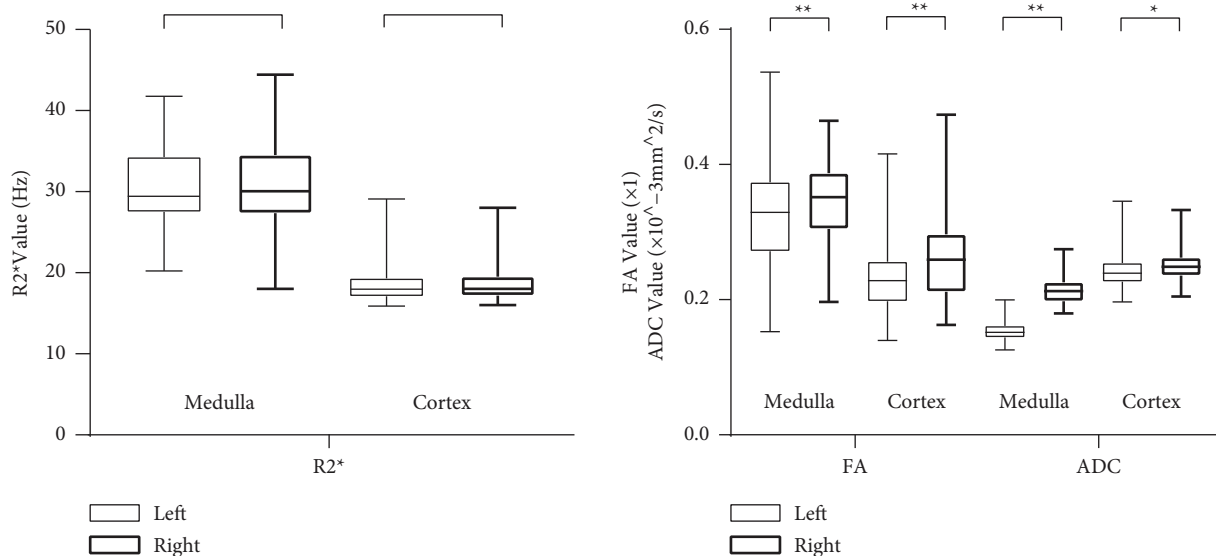
FIGURE 2: Comparison of the $R2^*$, ADC, and FA values between the right and left kidneys. Legends: Box plot comparing a total of six pairs including the $R2^*$ values, ADC values, and FA values between the left and right kidneys. $^*p < 0.01$, $^{**}p < 0.001$.

changes in renal oxygenation levels, water molecular diffusion, and renal microstructure during the progression of diabetes to DKD. Our results revealed that the $R2^*$ and FA values in the renal medulla were higher than those in the cortex in all the groups, and the same result was also seen in previous studies [21–24]. This demonstrated that the renal medulla was in a state of "physiological hypoxia" and provided evidence of a regular arrangement of radiating

medullary loops, collecting ducts, and tiny vessels. $R2^*$ and FA values in DKD patients showed significant differences from those in the HCs. The decline in FA was expected because changes in renal pathology, including glomerulosclerosis, interstitial fibrosis, tubular injury, and inflammatory cell infiltration, altered the anisotropy of well-aligned tubules [25, 26]. Notably, in our study, medullary $R2^*$ had a clear tendency to rise at the very early stage of

TABLE 4: Comparisons of $R2^*$ values, FA values, and ADC values of the cortex and medulla among HC and DM subgroups.

		HC	DI	DII	DIII	p	$p^{\#}$	$p^{\&}$	p^*
Cortex	$R2^*$ ms^{-1}	18.39 ± 0.82	19.41 ± 2.78	19.01 ± 2.03	17.50 ± 1.62	0.290	—	—	—
	FA	0.244 ± 0.062	0.231 ± 0.052	0.244 ± 0.068	0.266 ± 0.074	0.274	—	—	—
	ADC (×10^{-3} mm^2/s)	0.255 ± 0.032	0.241 ± 0.022	0.244 ± 0.024	0.246 ± 0.036	0.583	—	—	—
Medulla	$R2^*$ ms^{-1}	31.07 ± 2.11	33.12 ± 3.66	30.92 ± 4.46	25.46 ± 3.43	0.001	—	0.003	0.011
	FA	0.386 ± 0.080	0.322 ± 0.056	0.291 ± 0.055	0.280 ± 0.069	0.002	0.005	0.005	—
	ADC (×10^{-3} mm^2/s)	0.205 ± 0.022	0.204 ± 0.023	0.202 ± 0.016	0.217 ± 0.027	0.498	—	—	—

Legends: The results are expressed as the mean ± SD. $R2^*$, rate of spin dephasing (Hz); FA, fractional anisotropy; ADC, average diffusion coefficient (×10^{-3} mm^2/s). p: ANOVA was used to compare the values among all above groups, and a p value < 0.05 was considered significant. $p^{\#}$: Post-hoc paired comparisons between the HC group and the DII group; $p^{\&}$: Post-hoc paired comparisons between the HC group and the DIII group; p^*: Post-hoc paired comparisons between the DI and the DIII group.
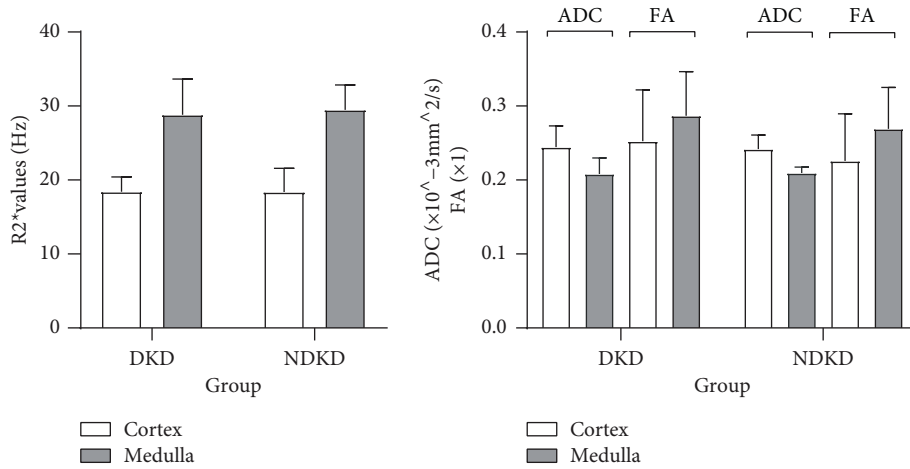
FIGURE 3: Comparisons of the $R2^*$ values, FA values, and ADC values of the cortex and medulla between the DKD and NDKD groups. Legends: Interleaved bars comparing the $R2^*$ values, FA values, and ADC values of the cortex and medulla between the DKD and NDKD groups. A p value < 0.05 was considered significant. No differences were found in the $R2^*$ values, FA values, or ADC values of the cortex and medulla between the DKD and NDKD groups.

DKD and rapidly decline once renal function was significantly impaired, which was same with the previous studies [27, 28].

The oxygenation level of the human kidney depends on the careful and coordinated physiological structure, and its coordination is mainly reflected in the balance between the oxygen supply and oxygen consumption of the kidneys. From a hemodynamic point of view, blood flow and the oxygen-carrying capacity of arterial blood are mainly responsible for the activities of supplying oxygen to the kidneys. Processes such as renal tubular reabsorption also utilize the oxygen transported to the kidneys. The blood flow of the renal cortex is significantly higher than the blood flow of the renal medulla, and the renal medulla is in a "physiological hypoxia state". The difference in blood flow between the cortex and medulla helps maintain the osmotic pressure gradient, which is beneficial for the body to concentrate urine. The progression of kidney disease often changes in function earlier instead of that in structure. The level of renal oxygenation is a very potential and particularly suitable indicator for the early evaluation of DKD. At present, chronic hypoxia has been found in animal models of DKD, and its possible mechanism is related to factors such as hypoxia-inducible factors, cytokines and inflammatory mediators, and renal tubular epithelial cell trans differentiation. In animal models, improving chronic hypoxia can indeed delay the progression of chronic kidney disease to a certain extent [11].

BOLD-MRI, using deoxyhemoglobin as an endogenous contrast agent, relies on its paramagnetic properties to obtain images that are sensitive to local tissue oxygen concentration. The increase of deoxyhemoglobin content in the blood causes the inhomogeneity of the magnetic field in the surrounding local tissue, which leads to the rapid dephasing of protons in the tissue and thus the signal decreases on $T2^*WI$ images. According to formulas (1) and (2),

$$SI(t) = SI_0 e^{-t/qT2^*}, \tag{1}$$

$$R2^* = 1/T2^*. \tag{2}$$

The apparent relaxation rate $R2^*$ can be obtained by calculating the monoexponential curve fitting of the signal intensity of a series of $T2^*WI$ images at different echo times to the TE time. The ratio of oxyhemoglobin to deoxyhemoglobin is related to the partial pressure of oxygen in the blood. The partial pressure of oxygen in the capillaries being in equilibrium with the surrounding tissue, the signal
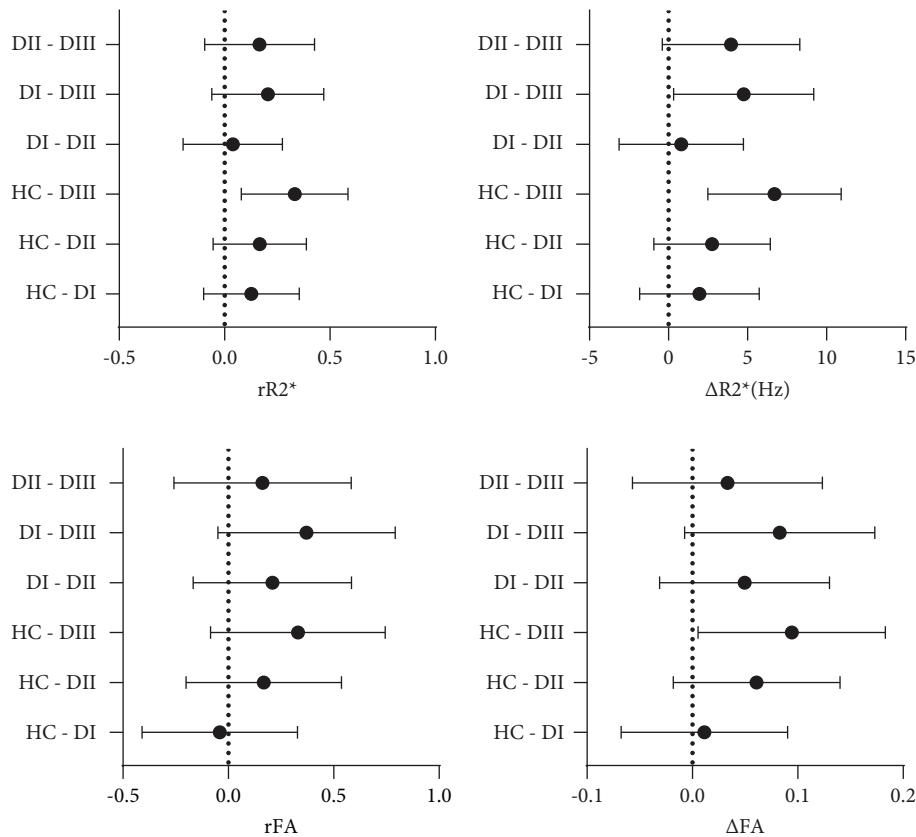
FIGURE 4: The D value and MCR evaluating the differences between the renal medulla and cortex among the groups. Legends: Plots of the 95% confidence intervals analyzing the $R2^*$ and FA values through D-values (e.g., $\Delta R2^* = MR2^* - CR2^*$) and MCRs (e.g., $rR2^* = MR2^*/CR2^*$) of the renal medulla and cortex among the groups. The range of values that did not fall on the dotted line was considered to be significantly different.

changes detected by BOLD-MRI can be interpreted as changes in the partial pressure of oxygen in the tissue. The decrease of $R2^*$ value indicates that the concentration of oxygenated hemoglobin increases in the local tissue, the proportion of deoxygenated hemoglobin decreases, and the local partial pressure of PaO_2 increases. BOLD-MRI technology was first applied to the study of brain function, and its basis mainly depends on the close connection between hemodynamics and brain neural activity. Prasad et al. were the first scholars to apply BOLD-MRI to assess the level of renal oxygenation in humans. They also found that the medullary $R2^*$ value of patients with kidney disease showed a significant decreasing trend after using diuretics, suggesting that diuretics can improve renal hypoxia [29]. In this study, for the subjects recruited, we excluded those who used diuretics and those taking antihypertensive drugs containing diuretics in the past month to prevent interference errors to $R2^*$ values.

$R2^*$ is proportional to the content of deoxyhemoglobin, so an increase in $R2^*$ represents a decrease in PO_2 and tissue hypoxia [30]. DM patients might experience renal medulla hypoxia, which is related to the highly increased oxygen consumption and less blood supply in the medulla, such as active reabsorption of excess sodium, glomerular hyperfiltration, and increased Na+/K+–ATPase activity in a hyperglycemic state [28, 31]. Significant elevation of $MR2^*$ was shown in DM patients without kidney disfunction in our study. Considering that these patients had higher HGB values and eGFR levels compared with healthy controls and DKD patients, the elevated eGFR and HGB seemed to verify that the kidney tissue was in a state of hypoxia compensation. At this stage, some diabetic patients may experience increased blood pressure, and it is recommended that blood pressure should not be controlled too low. The removal of the high filtration state will cause insufficient renal blood perfusion and accelerate renal damage [32]. Besides, $R2^*$ was also affected by other factors, such as hemodynamic and structural changes, red blood cell volume and the vascular volume fraction, vessel geometry, and applied pulse sequence parameters, including factors influencing the oxygen dissociation curve [33]. In the later stages of DKD, hypoxia indeed takes place in the medulla despite the decreased $R2^*$ in diabetic patients with a moderately decreased eGFR. It was inferred that the eGFR decreased and the tubular reabsorption decreased, along with the metabolic products of renal tubular epithelial cells, interstitial inflammatory cells, and endothelial cells that stimulate interstitial inflammation and fibrosis in the kidney, so that oxygen consumption was decreased [34]. The phenomenon of "hypoxic remission" of the medulla appeared. Additionally, there was no significant
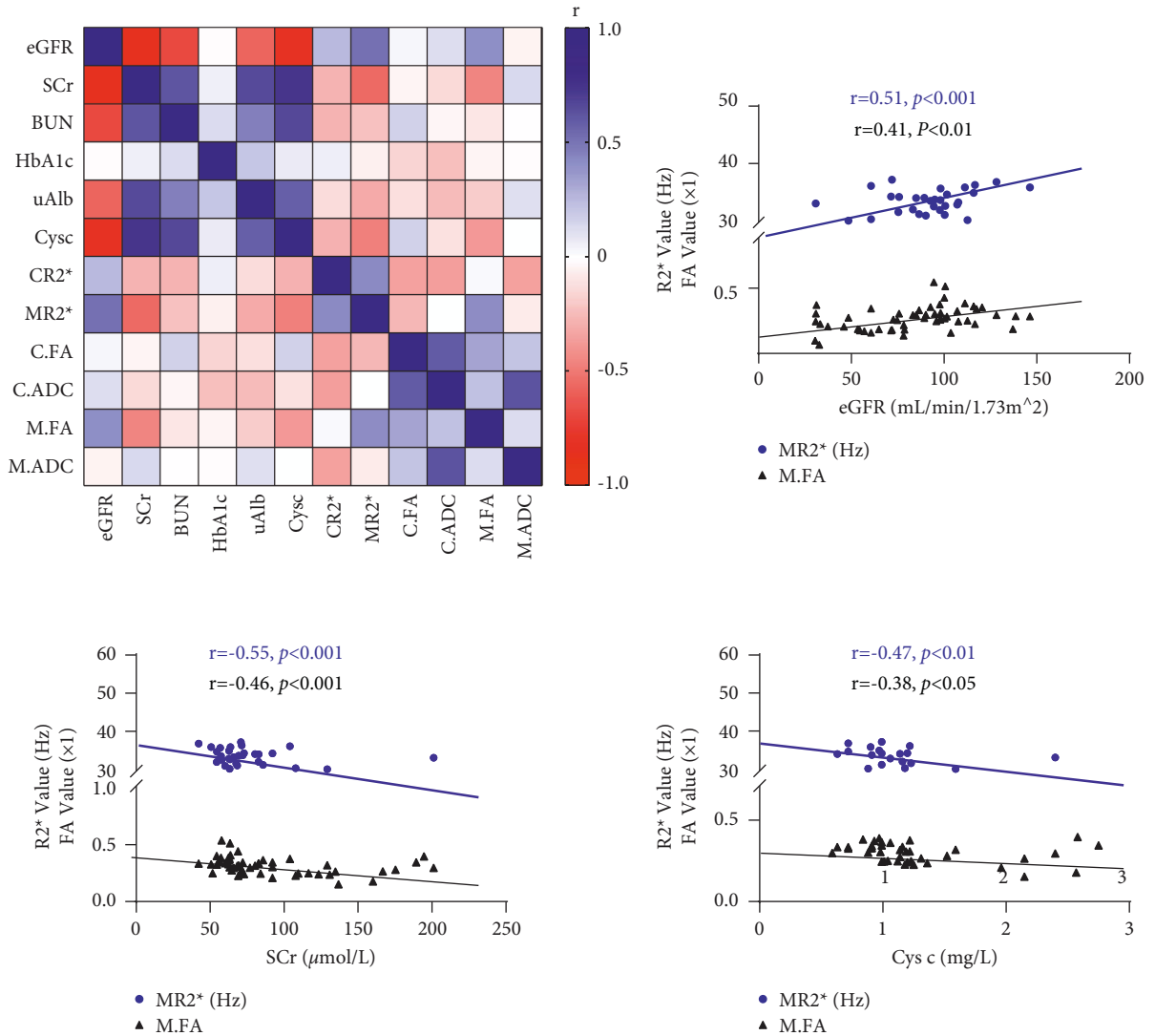
FIGURE 5: Correlations between clinical indexes and MRI parameters. Legends: Heatmap analyzing the correlation between the cortex and medulla $R2^*$, FA, and ADC values and the eGFR, BUN, HbA1c, uAlb, and Cys C levels of all the groups. The colormap was drawn with a double gradient. The r value ranges from −1 to 1. Blue represents the maximum value and red represents the minimum value. The linear correlations of clinical indexes with MRI data $MR2^*$/M.FA were also shown above. A p value < 0.05 was considered significant.

difference on $MR2^*$ between the DKD and NDKD, implying that diabetic patients with different etiologies of CKD were consistent in their changes in renal oxygenation levels.

DWI imaging can noninvasively evaluate the movement of water molecules in vivo. In recent years, the study of renal DWI under physiological or pathological conditions has gradually become a hot topic. Diffusion tensor imaging (DTI-MRI) is a technology developed based on the DWI technology. It was an MRI technique proposed in 1994 to analyze white matter fiber tracts. It reflects the tensor information of the diffusional motion of water molecules in vivo by applying diffusion-sensitive gradients in at least 6 noncollinear directions. So, the description of water molecular motion is more precise. DTI employs parameter values to reflect tissue diffusion characteristics. An FA value, ranging from 0 to 1, measures the proportion of the anisotropic component of the diffusion of water molecules in the overall diffusion tensor. The DTI postprocessing

program of the workstation can correct the EPI distortion. The formula (3) required to generate the FA map is as follows. λ_1, λ_2, and λ_3 are the eigenvalues of the diffusion tensor.

$$FA = \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2}{2(\lambda_1 + \lambda_2 + \lambda_3)^2}}. \tag{3}$$

The larger the value, the more restrictive the tissue is to the direction of diffusion of water molecules. Compared with DWI, DTI-MRI can not only observe the restricted degree of the diffusion movement of water molecules but also show the differences in the directionality of the diffusion movement of water molecules and provide information about the changes in the ultrastructure of the tissue, helping evaluate the subtleties in the structure of the tissue. This makes it possible to apply DTI imaging to the evaluation of kidney-related physiological

function and pathological states. Specifically, DTI can indirectly reflect microstructural changes of normal or damaged tissues by diffusion-related parameters, such as ADC and FA values. ADC focuses on the intensity or freedom, and FA focuses on the direction of water molecular diffusion. With respiratory triggering technology, the image quality of renal DTI is maintained better despite the influence of other soft tissue and the breathing movement of the native kidney. Currently, it can be widely used in the study of autologous kidneys and transplanted kidneys.

In general, cortical blood flow was abundant, and the radially arranged tubule structure in the medulla limited the diffusion of water molecules, so the ADC value was higher and the FA value was lower in the cortex [35]. In our study, the cortical and medullary ADCs did not differ significantly between HCs and the three diabetic subgroups. But diabetic patients had slightly lower cortical ADCs than compared with HCs, and DKDs with moderately decreased eGFR had higher medullary ADCs than the other groups. The reason relied on the possibility that diabetic patients had higher blood viscosity than HCs, and the disruption of the initially regularly arranged microstructure in the renal medulla could no longer well restrict the free diffusion of water molecules. Medullary FA decreased gradually with the decline in eGFR in diabetic patients since the kidneys in early-stage DKDs were likely to undergo microscopic changes such as tubular ectasia, tubular epithelial cell edema and necrosis, and interstitial inflammatory cell infiltration and proliferation, which destroyed the microstructures and inhibited the movement of water molecules [36, 37]. The ADC and FA results in Feng's DTI study were similar to our study. It seemed that FA had more potential than ADC in determining the diffusion of water molecules and the microstructural changes in the human kidneys [33]. However, in some DWI studies, the advantages of apparent diffusion coefficient (ADC) values were reflected. Wang et al. found that the cortical ADC value in the DKD murine model was higher than in the blank model [38]. Emre et al. revealed that ADC values of renal parenchyma could be used for early diagnosis and clinical staging of patients with kidney disease [39]. Xu et al. found a negative correlation between renal parenchymal ADC values and the degree of interstitial fibrosis [40].

Medullary $R2^*$ and FA had positive correlations with eGFR and negative correlations with SCr and Cys C, which is consistent with Lu L's study [41]. Serum Cys C is filtered through the glomerulus, entirely reabsorbed, catabolized, and metabolized by the proximal tubular epithelium. In addition, Cys C is not disturbed by inflammatory reactions, tumors, hemolysis, bilirubin, triglycerides, etc., and is also independent of sex, age, muscle mass, and diet. Cys C is a more sensitive indicator than SCr for the diagnosis of renal impairment. Mojiminyi et al. [42] suggested that Cys C is a more sensitive and practical indicator for detecting DKD.

The advantages of BOLD and DTI allow them to refrain from radiation and there is no need for an injection of a contrast agent. The renal cortex and medulla worked together and somewhat differently and maintained a clear functional and structural differentiation. D-value and MCR compared the difference between the cortex and medulla in an individual kidney rather than focusing only on the changes in medullary $R2^*$ and FA values. Compared with HCs and simple DM patients, statistically reduced cortical-medullary differentiation of the kidney was shown in DKD patients with microalbuminuria and massive albuminuria. With the improvement of technical methods and the continuous optimization of postprocessing analysis, DTI will surely show more special applications in scientific research and clinical practice. Its prospects are very broad, especially combined with IVIM technology and BOLD-fMRI in the field of nephrology.

4.1. Limitations of the Study. The sample size of this study is relatively small. The number of cases in the subgroups requires amplification. In addition, this study only investigated BOLD and DTI parameters in assessing the changes in renal oxygenation levels and microstructure. It did not fully elucidate the mechanisms of renal hypoxia or provide insight into the specific changes in the structure in DKD combined with the pathological basis. Alterations in microcirculatory perfusion could also affect the result. At present, there is no unified technical standard for BOLD-MRI to detect renal oxygenation at home or abroad. The normal value of the relaxation rate is related to different field strengths. In DTI, diffusion gradients could cause eddy currents, which have an effect on the final parameters. Combining with arterial spin labeling (ASL) and IVIM, BOLD and DTI may be useful for a more comprehensive assessment of renal functional status.

5. Conclusion

In conclusion, our study shows that BOLD and DTI are effective and safe ways to estimate renal function through the oxygenation level, water molecular movement, and microstructural alterations in patients with DKD and it proves that BOLD performs better than DTI. With renal function declines, alterations can be found including the firstly rising and then falling medullary $R2^*$, continuously declining medullary FA and apparent cortex-medullary differentiation in DKD patients.

Abbreviations

ADA: American Diabetes Association
BUN: Blood urea nitrogen
BMI: Body mass index
C.ADC: Cortical average diffusion coefficient
$CR2^*$: Cortical $R2^*$
Cys C: Cystatin C
DM: Diabetes mellitus
DKD: Diabetic kidney disease
eGFR: Estimated glomerular filtration rate
EPI: Echo-planar imaging
HbA1c: Hemoglobin A1
HC: Healthy control
HGB: Hemoglobin
ICC: Intraclass correlation analysis
KDIGO: Kidney Disease: Improving Global Outcomes

M.ADC: Medullary average diffusion coefficient
MDRD: Modification of Diet in Renal Disease
M.FA: Medullary FA
MFGRE: Multiple fast gradient recalled echo
$MR2^*$:   Medullary $R2^*$
NDKD: Nondiabetic kidney disease
ROI: Regions of interest
SCr: Serum creatinine
SE: Spin-echo
UACR: Urinary albumin-creatinine ratio
uAlb: Urinary microalbumin.

Data Availability

The data used to support the findings of this study are restricted by the data providers.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Authors' Contributions

Xiaobao Wei and Runyue Hu contributed equally to this work. Manuscript review and final approval were conducted by Xiaoyan Wu, Haibo Xu.

References

[1] K. C. Mekala and A. G. Bertoni, "Epidemiology of diabetes mellitus," in *Transplantation, Bioengineering, and Regeneration of the Endocrine Pancreas*, G. Orlando, Ed., pp. p49–58, Academic Press, 2020.

[2] M. Afkarian, L. R. Zelnick, Y. N. Hall et al., "Clinical manifestations of kidney disease among US adults with diabetes, 1988-2014," *JAMA*, vol. 316, no. 6, pp. 602–610, 2016.

[3] I. H. de Boer, "Kidney disease and related findings in the diabetes control and complications trial/epidemiology of diabetes interventions and complications study," *Diabetes Care*, vol. 37, no. 1, pp. 24–30, 2014.

[4] A. O. Luk, E. M. Hui, M. C. Sin et al., "Declining trends of cardiovascular-renal complications and mortality in type 2 diabetes: the Hong Kong Diabetes Database," *Diabetes Care*, vol. 40, no. 7, pp. 928–935, 2017.

[5] L. Zhang, J. Long, W. Jiang et al., "Trends in chronic kidney disease in China," *New England Journal of Medicine*, vol. 375, no. 9, pp. 905-906, 2016.

[6] K. R. Tuttle, G. L. Bakris, R. W. Bilous et al., "Diabetic kidney disease: a report from an ADA Consensus Conference," *Diabetes Care*, vol. 37, no. 10, pp. 2864–2883, 2014.

[7] I. H. De Boer, M. L. Caramori, J. C. N. Chan et al., "Kidney disease: improving global outcomes (KDIGO) diabetes work group KDIGO 2020 clinical practice guideline for diabetes management in chronic kidney disease," *Kidney International*, vol. 98, no. 4, pp. S1–S115, 2020.

[8] R. Retnakaran, C. A. Cull, K. I. Thorne, A. I. Adler, and R. R. Holman, "Risk factors for renal dysfunction in type 2 diabetes: UK Prospective Diabetes Study 74," *Diabetes*, vol. 55, no. 6, pp. 1832–1839, 2006.

[9] A. J. W. Branten, G. Vervoort, and J. F. M. Wetzels, "Serum creatinine is a poor marker of GFR in nephrotic syndrome," *Nephrology Dialysis Transplantation*, vol. 20, no. 4, pp. 707–711, 2005.

[10] M. Takemoto, M. Shoji, and K. Yokote, "Urinary podocyte and urinary podocalyxin; markers for diabetic nephropathy," *Nihon rinsho. Japanese journal of clinical medicine Japanese journal of clinical medicine*, vol. 74, no. 2, pp. 181–185, 2016.

[11] K. U. Eckardt, W. W. Bernhardt, A. Weidemann et al., "Role of hypoxia in the pathogenesis of renal disease," *Kidney International*, vol. 68, pp. S46–S51, 2005.

[12] M. Friederich-Persson, E. Thörn, P. Hansell, M. Nangaku, M. Levin, and F. Palm, "Kidney Hypoxia, Attributable to increased oxygen consumption, induces nephropathy independently of hyperglycemia and oxidative stress," *Hypertension*, vol. 62, no. 5, pp. 914–919, 2013.

[13] T. A. Schiffer and M. Friederich-Persson, "Mitochondrial reactive oxygen species and kidney hypoxia in the development of diabetic nephropathy," *Frontiers in Physiology*, vol. 8, no. 8, p. 211, 2017.

[14] W. H. Yiu, K. W. Chan, L. Y. Y. Chan et al., "Complement C5a inhibition moderates lipid metabolism and reduces tubulointerstitial fibrosis in diabetic nephropathy," *Nephrology Dialysis Transplantation*, vol. 33, no. 8, pp. 1323–1332, 2018.

[15] S. N. Heyman, M. Khamaisi, S. Rosen, and C. Rosenberger, "Renal parenchymal hypoxia, hypoxia response and the progression of chronic kidney disease," *American Journal of Nephrology*, vol. 28, no. 6, pp. 998–1006, 2008.

[16] R. W. Chan, C. Von Deuster, C. T. Stoeck et al., "High-resolution diffusion tensor imaging of the human kidneys using a free-breathing, multi-slice, targeted field of view approach," *NMR in Biomedicine*, vol. 27, no. 11, pp. 1300–1312, 2014.

[17] K. Hueper, D. Hartung, M. Gutberlet et al., "Magnetic resonance diffusion tensor imaging for evaluation of histopathological changes in a rat model of diabetic nephropathy," *Investigative Radiology*, vol. 47, no. 7, pp. 430–437, 2012.

[18] American Diabetes Association, "11. Microvascular complications and foot care: *Standards of m*," *Diabetes Care*, vol. 43, no. Supplement_1, pp. S135–S151, 2020.

[19] R. Katafuchi, Y. Kiyoshi, Y. Oh et al., "Glomerular score as a prognosticator in IgA nephropathy: its usefulness and limitation," *Clinical Nephrology*, vol. 49, no. 1, pp. 1–8, 1998.

[20] P. Soininen, A. J. Kangas, P. Wurtz, T. Suna, and M. Ala-Korpela, "Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics," *Circulation Cardiovascular genetics*, vol. 8, no. 1, pp. 192–206, 2015.

[21] M. L. Gloviczki, J. Glockner, S. I. Gomez et al., "Comparison of 1.5 and 3 T BOLD MR to study oxygenation of kidney cortex and medulla in human renovascular disease," *Investigative Radiology*, vol. 44, no. 9, pp. 566–572, 2009.

[22] M. Cutajar, J. D. Clayden, C. A. Clark, and I. Gordon, "Test-retest reliability and repeatability of renal diffusion tensor MRI in healthy subjects," *European Journal of Radiology*, vol. 80, no. 3, pp. e263–8, 2011.

[23] A. A. K. A. Razek, M. A. A. A. Al-Adlany, A. M. Alhadidy, M. A. Atwa, and N. E. A. Abdou, "Diffusion tensor imaging of the renal cortex in diabetic patients: correlation with urinary and serum biomarkers," *Abdom Radiol*, vol. 42, no. 5, pp. 1493–1500, 2017.

[24] M. Ries, R. A. Jones, F. Basseau, C. T. Moonen, and N. Grenier, "Diffusion tensor MRI of the human kidney," *Journal of Magnetic Resonance Imaging*, vol. 14, no. 1, pp. 42–49, 2001.

[25] J. Y. Kaimori, Y. Isaka, M. Hatanaka et al., "Visualization of kidney fibrosis in diabetic nephropathy by long diffusion tensor imaging MRI with spin-echo sequence," *Scientific Reports*, vol. 7, no. 1, p. 5731, 2017.

[26] Y. Y. Yan, S. Hartono, T. Hennedige et al., "Intravoxel incoherent motion and diffusion tensor imaging of early renal fibrosis induced in a murine model of streptozotocin induced diabetes," *Magnetic Resonance Imaging*, vol. 38, no. Complete, pp. 71–76, 2017.

[27] P. Prasad, L. P. Li, and S. Halter, "Evaluation of renal hypoxia in diabetic mice byBOLD," *BOLD MRI. Invest Radiol*, vol. 45, no. 12, pp. 001–004, 2010.

[28] W. J. Yin, F. Liu, X. M. Li et al., "Noninvasive evaluation of renal oxygenation in diabetic nephropathy by BOLD-MRI," *European Journal of Radiology*, vol. 81, no. 7, pp. 1426–1431, 2012.

[29] P. V. Prasad, R. R. Edelman, and F. H. Epstein, "Noninvasive evaluation of intrarenal oxygenation with BOLD MRI," *Circulation*, vol. 94, no. 12, pp. 3271–3275, 1996.

[30] S. M. Tumkur, A. T. Vu, L. P. Li, L. Pierchala, and P. V. Prasad, "Evaluation of intra-renal oxygenation during water diuresis: a time-resolved study using BOLD MRI," *Kidney International*, vol. 70, no. 1, pp. 139–143, 2006.

[31] E. A. dos-Santos, L. P. Li, L. Ji, and P. V. Prasad, "Early changes with diabetes in renal medullary hemodynamics as evaluated by fiberoptic probes and BOLD magnetic resonance imaging," *Investigative Radiology*, vol. 42, no. 3, pp. 157–162, 2007.

[32] Y. Z. Feng, Y. J. Ye, Z. Y. Cheng et al., "Non-invasive assessment of early stage diabetic nephropathy by DTI and BOLD MRI," *British Journal of Radiology*, vol. 93, no. 1105, p. 20190562, 2020.

[33] T. A. Schiffer and M. Friederich-Persson, "Mitochondrial reactive oxygen species and kidney hypoxia in the development of diabetic nephropathy," *Frontiers in Physiology*, vol. 8, p. 211, 2017.

[34] Z. L. Li and B. C. Liu, "Hypoxia and renal tubulointerstitial fibrosis," *Advances in Experimental Medicine & Biology*, vol. 1165, pp. 467–485, 2019.

[35] H. C. Thoeny, F. De Keyzer, R. H. Oyen, and R. R. Peeters, "Diffusion-weighted MR imaging of kidneys in healthy volunteers and patients with parenchymal diseases: initial experience," *Radiology*, vol. 235, no. 3, pp. 911–917, 2005.

[36] Q. Feng, Z. Ma, J. Wu, and W. Fang, "DTI for the assessment of disease stage in patients with glomerulonephritis- correlation with renal histology," *European Radiology*, vol. 25, no. 1, pp. 92–98, 2015.

[37] L. Xie, R. Dibb, G. P. Cofer et al., "Susceptibility tensor imaging of the kidney and its microstructural underpinnings," *Magnetic Resonance in Medicine*, vol. 73, no. 3, pp. 1270–1281, 2015.

[38] W. Yue and Y. Wu, "A preliminary study of the therapeutic effect of traditional Chinese medicine Tangshen Formula on early diabetic nephropathy evaluated by BOLD -MRI and DTI," *Chinese Journal of Integrative Imaging*, vol. 19, no. 3, p. 65, May, 2021.

[39] T. Emre, O. Kilickesmez, A. Buker, B. B. Inal, H. Dogan, and T. Ecder, "Renal function and diffusion-weighted imaging: a new method to diagnose kidney failure before losing half function," *Radiologia Medica, La*, vol. 121, no. 3, pp. 163–172, 2016.

[40] X. Xu, S. L. Palmer, X. Lin et al., "Diffusion-weighted imaging and pathology of chronic kidney disease: initial study," *Abdom Radiol (NY)*, vol. 43, no. 7, pp. 1749–1755, 2018.

[41] L. Lu, J. R. Sedor, V. Gulani et al., "Use of diffusion tensor MRI to identify early changes in diabetic nephropathy," *American Journal of Nephrology*, vol. 34, no. 5, pp. 476–482, 2011.

[42] O. A. Mojiminiyi, N. Abdella, and S. Grorge, "Evaluation of serum Cyatatin C and chromogranin A as markers of nephropathy in patients with typediabetes mellitus," *Scandinavian Journal of Clinical & Laboratory Investigation*, vol. 60, no. 6, pp. L483–L489, 2000.

Research Article

Multivariable Linear Algebraic Discretization of Nonlinear Parabolic Equations for Computational Analysis

Li Zuo [1] and **Fengtai Mei** [2]

[1]*School of Nursing, Chengdu Polytechnic, Chengdu 610071, China*
[2]*College of General Education, Chengdu Polytechnic, Chengdu 610071, China*

Correspondence should be addressed to Li Zuo; zuoli@mjc-edu.cn

Since the nonlinear parabolic equation has many variables, its calculation process is mostly an algebraic operation, which makes it difficult to express the discrete process concisely, which makes it difficult to effectively solve the two grid algorithm problems and the convergence problem of reaction diffusion. The extended mixed finite element method is a common method for solving reaction-diffusion equations. By introducing intermediate variables, the discretized algebraic equations have great nonlinearity. To address this issue, the paper proposes a multivariable linear algebraic discretization method for NPEs. First, the NPE is discretized, the algebraic form of the nonlinear equation is transformed into the vector form, and the rough set (RS) and information entropy (IE) are constructed to allocate the weights of different variable attributes. According to the given variable attribute weight, the multiple variables in the equation are discretized by linear algebra. It can effectively solve the two grid algorithm problems and the convergence problem of reaction diffusion and has good adaptability in this field.

1. Introduction

When the relationship between the two variables is not linear, it is said to be nonlinear parabolic. Some of the relations could be square, logarithmic, exponential, trigonometric function relationships, etc. Because a parabolic equation can describe the state or process of a physical quantity changing with time, it has many applications in real life, such as heat conduction, liquid permeation, and gas permeation. For example, the heat conduction NPE is a differential equation that describes the dispersion of atmospheric pollutant concentration, coastal salinity, and the law of fluid motion, while the delay NPE is also widely used, such as in the fields of population dynamics, ecology, and environmental science. The data simulation problems can be reduced to a delay NPE [1]. There are many variables in the NPE. When the NPE is used in practice, the amount of data corresponding to each variable is huge. When the number of variables in the NPE increases, it becomes very difficult to solve the equation. Therefore, it is necessary to discretize the variables of NPE.

At present, some scholars have studied it. For example, document [2] proposes a three-step, two-layer grid method for the nonlinear reaction-diffusion problem, that is, solving the original nonlinear algebraic equations on the coarse grid (CG) and then correcting the CG . Through convergence analysis, it is found that the three-step two-layer mesh method can maintain the asymptotic optimal approximation of the mixed finite element solution when the selected coarse space step size is satisfied. However, this method has high computational complexity and low efficiency. When the variable data distribution is non-Gaussian, the second-order statistics method that relies on the data is often ineffective. It does not explicitly give the number of principal components, which may affect the results of discretization. In addition, some traditional discretization methods are mostly aimed at a single variable, and the calculation process is mostly an algebraic operation, making it difficult to express the discrete process concisely [3]. Therefore, to solve the two grid algorithm problems of NPE and the convergence problem of reaction diffusion, this paper will study the multivariable

linear algebraic discretization method of NPE to provide some help for solving the multivariable linear algebraic discretization problem of NPE.

2. Multivariable Linear Algebraic Discrete Methods for NPEs

2.1. Discretization of NPEs. For the NPE, the discretization of variables is used to discretize the attributes of the variables to solve the NPE. Discretization is the process of mapping a finite number of individuals in an infinite space into a finite space to improve the spatial and temporal efficacy of data or variables. Discretization is the process of reducing the size of data without making changes in the relative size of the variable data. Discretization essentially comes down to the problem of dividing the space of conditional attributes by using the selected breakpoints. The space of attributes is divided into a finite area, so that the decision values of objects in each area are the same. If an attribute has an attribute value, then there is a breakpoint on the attribute that is desirable. As the number of attributes increases, the number of desirable breakpoints increases geometrically. The process of selecting breakpoints is also the process of merging attribute values [4]. By merging attribute values, the number of attribute values and the number of breakpoints can be reduced, and thus the complexity of the problem can be reduced. The current discretization methods are all aimed at single variables, and they consider only one attribute for discrete attributes. Therefore, the result of the discretization of a single variable is often not optimal because the target class in variable data is determined by multiple attributes rather than by a single attribute.

To discretize the NPE, the scheme of the NPE should be changed. The discrete expression of the NPE is as follows:

$$
\begin{cases}
u_t - u_{xx} - \dfrac{\sigma}{x} u = f(u),\, (x,t) \in \Omega = (0,1) \times (0,T), \\[2mm]
u_x(0,t) = u(1,t) = 0,\, t \in (0,T), \\[2mm]
u(x,0) = 0,\, x \in \Gamma, \\[2mm]
\Gamma = \{x : x \in (0,1)\}.
\end{cases}
\tag{1}
$$

In the above formula (1), the nonhomogeneous term $f(u)$ at the right end of the equation satisfies continuity, that is, for $\forall v, \omega \in R$, there is a constant L, which makes $|f(x,t,v) - f(x,t,\omega)| \le L|v - \omega|$ hold. Take the space step $h = 1/M$, the time step $k = 1/N$ (M, N are positive integers), and note that the space node is $x_m = mh$, the time node is $t_n = nk$, and the numerical solution is $un/m \approx u(x_m, t_n)$. If the solution u of formula (1) has the necessary differentiability, then it is as follows:

$$
u(x_{m+1}, t) - 2u(x_m, t) + u(x_{m-1}, t)
$$
$$
= \frac{\partial^2 u(x_m, t)}{\partial x^2} h^2 + \frac{1}{12} \frac{\partial^4 u(x_m, t)}{\partial x^4} h^4 + o(h^6).
\tag{2}
$$

After a series of simplification and consolidation operations on formula (2), according to the relationship existing in formula (1).

$$
\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{\partial u(x,t)}{\partial t} - f(x,t).
\tag{3}
$$

From the Taylor expansion (3), we can get the following difference transformation scheme for NPE.

$$
\frac{k}{h^2} \left[u(x_{m+1}, t) - 2u(x_m, t) \right] + u(x_{m-1}, t)
$$
$$
= \frac{k}{12} \frac{\partial u(x_{m+1}, t)}{\partial t} + \frac{10k}{12} \frac{\partial u(x_m, t)}{\partial t} + \frac{k}{12} \frac{\partial u(x_{m-1}, t)}{\partial t}
$$
$$
- \frac{k}{12} f(x_{m+1}, t) - \frac{10k}{12} f(x_m, t)
$$
$$
- \frac{k}{12} f(x_{m-1}, t) + o(k^2 + kh^4).
\tag{4}
$$

To make a linear algebraic discretization of the transformed NPE, the difference scheme shown in the above formula is transformed into the vector form shown in the following formula.

$$
rAu^n + Bu^n = Bu^{n-1} + kBf^{n-1} + d^n.
\tag{5}
$$

In the above formula (5), r, A, and B are all tridiagonal matrices of order $(M-1) \times (M-1)$. u^n, f^{n-1}, and d^n are all $M-1$-dimensional column vectors, and k is the Taylor expansion order. The parameters of matrix form of NPE are as follows:

$$
\begin{cases}
r = \dfrac{k}{h^2}, \\[3mm]
A = \operatorname{tri\,di\,ag}(-1, 2, -1), \\[3mm]
B = \operatorname{tir\,di\,ag}\left(\dfrac{1}{12}, \dfrac{10}{12}, \dfrac{1}{12}\right).
\end{cases}
\tag{6}
$$

After transforming the NPE into vector form, several variables in the equation are treated with linear algebraic discretization [5–7]. The RS and IE of NPE variables are constructed.

2.2. Construction of RSs and IE. To deal with imprecise data, RS is used. RSs do not need any prior knowledge or additional information about the data. In RS, a variable data table is called an information system (IS). The IS is a quad $G_{IS} = (U, W, V, f)$, which meets the following four conditions:

(1) U represents a nonempty set of objects.

(2) W represents a nonempty finite attribute set;

(3) V is the set of attribute values, $V = U(V_w)$, V_w is the value field of attribute w.

(4) $f: U \times W \longrightarrow V$ is a mapping function, which represents a value of the attribute value set mapped by each attribute of each object;

If an attribute from the attribute set is considered a decision attribute, IS G_{IS} is also called decision-table. Expressed as: $T_D = (U, C, D, V, f)$, among them, $C \cup D = W, C \cap D = \phi$; C is called the set of conditional attributes; D is the set of decision-attributes. In a given decision-table (DT), there is an indistinguishable relationship; that is, there is a certain interaction between the two variables. Indistinguishable relations are defined as follows:

for a given DT $T_D = (U, C, D, V, f)$ and $Q \in C \cup D$, the binary relation $IND(Q)$ is called indistinguishable relation. The definition formula is as follows:

$$IND(Q) = \{(x, y) \in U \times U\}. \tag{7}$$

The above formula (7) satisfies $\forall a \in Q(f(x, a) = f(y, a))$. Indistinguishable relation $IND(Q)$ is equivalent relation on U, which divides U into disjoint equivalent classes. $U/IND(Q)$ represents the set of equivalence classes of an equivalence relation $IND(Q)$ on U, U/Q for each object $x \in U$, the equivalent class of element x in U/Q is represented by $[x]Q$. It is called the equivalence class of object x on equivalence relation $IND(Q)$. Using the new DT instead of the original one can reduce the variable attributes of the NPE. After constructing the RS, the IE of the NPE is calculated.

IE is used to describe the purity of data sets. When data sets belong to a certain category, the IE is 0. When the data of data sets are more mixed, the IE is higher. IE and discrete breakpoint IE are defined as follows:

$$\begin{cases} E(S) = -\sum_{i=1}^{s} P(C_i, S)\log\left(P(C_i, S)\right), \\ \\ E(W, T, S) = \dfrac{|S_1|}{|S|} Ent(S_1) + \dfrac{|S_2|}{|S|} Ent(S_2). \end{cases} \tag{8}$$

In the above formula (8), S is the set of objects; s is the number of variables in the NPE; C_i represents the number of variables of type i in object set S; W, T represent breakpoint T on attribute W, respectively. $|S|$ is the cardinality of set S [8–10]. After constructing the RS and IE of the NPE, the breakpoint is selected and the discrete decision tree is established.

2.3. Model Problem.
Through the reaction diffusion problem of the mathematical model of the porous medium groundwater flow problem, and explain the physical meaning of the parameters in the model [11–13]. This model can be described by a set of nonlinear partial differential equations. The nonlinear reaction diffusion equation is given as follows:

$$\frac{\partial p}{\partial t} - \nabla \times (K(P)\nabla p) = f(p, \nabla p), (x, t) \in \Omega \times J. \tag{9}$$

The initial conditions are given as follows:

$$p(x, 0) = p^0(x), x \in \Omega. \tag{10}$$

The boundary conditions are given as follows:

$$K(P)\nabla p \times v = 0, (x, t) \in \partial\Omega \times J, \tag{11}$$

where $\Omega \subset R^2$, is the polygonal region whose boundary is marked as $\partial\Omega$ [14, 15]. Where v is the normal vector outside the unit of $\partial\Omega$, $J = (0, T]$ and K are the tensor of the square integrable symmetric positive definite. It is composed of the first order conservation of mass and the following equations with respect to energy P and relative velocity u.

$$\frac{\partial p}{\partial t} + \nabla \times u = f(p, \nabla p), \tag{12}$$

$$K(P)^{-1} u + \nabla p = 0, \tag{13}$$

where the strain in the equations (12) and (13) is p and u. p is the unknown fluid pressure; u is the flow rate of the liquid; and $f(p, \nabla p)$ is the external flow rate [16–18]. Firstly, the weak form of the equation and the fully discrete scheme of the extended mixed finite element are established. Then, the error estimates of the fully discrete scheme are obtained by using the projection operator and its approximation properties.

Let the following three variables, pressure p, gradient $\tau = \nabla p$, and flow $\varphi = K(p)\tau$. The weak form of the initial boundary value problem is defined, i e. there is $(p, \tau, \varphi) \in W \times V \times \widetilde{V}$, so that,

$$\left(\frac{\partial p}{\partial t}, \omega\right) + (\nabla \times \varphi, \omega) = (f(p, \tau), \omega), \omega \in W, \tag{14}$$

$$(\tau, v) = -(p, \nabla \times v), v \in \widetilde{V}, \tag{15}$$

$$(\varphi, v) = (K(p)\tau, v), v \in V, \tag{16}$$

according to the weak form (14)–(16) of the initial boundary value problem, RT_k group of time discrete extended mixed finite element approximations can be established on the a-element, given $(pn/h, \tau n/h, \varphi n/h) \in W_h \times V_h \times \widetilde{V}_h$, for $n, n = 1, \Lambda, N$, let $(pn/h, \tau n/h, \varphi n/h) \in W_h \times V_h \times \widetilde{V}_h$ satisfy the following formula:

$$\left(\frac{p_h^n - p_h^{n-1}}{t}, \omega_h\right) + (\nabla \times \varphi_h^n, \omega_h) = (f(p_h^n, \tau_h^n)), \omega_h \in W_h, \tag{17}$$

$$(\tau_h^n, v_h^n) = -(p_h^n, \nabla \times v_h), v_h \in \widetilde{V}_h, \tag{18}$$

$$(\varphi_h^n, v_h) = (K(p_h^n)\tau_h^n, v_h), v_h \in V_h. \tag{19}$$

The existence and uniqueness of the solutions of the nonlinear equations (17)–(19) have been proved.

At the same time, the minimum number of breakpoints is obtained and the indiscernible relationship between objects is ensured. The reasonable criteria for breakpoint selection are generally: consistency, irreducibility, and minimum discreteness. Given a DT $T_D = (U, C, D, V, f)$,

and C_i is the candidate breakpoint (CBP) set of the i conditional attribute $a_i \in C$ on the domain U. C_s is a subset of U, C_{si} is the CBP set of the i-th conditional attribute $a_i \in C$ on the domain C_s, where C_{si} is a subset of C_i.

In this paper, the label $\{a_i \,|\, C_s\}$ is used to represent the set of value fields of attribute a_i on a subdomain U_s. The formalized formula is as follows:

$$C_{si} = \left\{ c \,\middle|\, c \in C_i \wedge \min\left(a_i \,\middle|\, C_s\right) < c < \max\left(a_i \,\middle|\, C_s\right) \right\}. \quad (20)$$

The specific treatment process is given as follows:

(1) If the DT contains numerical attributes, the algorithm afc4.5 is used to discretize them.

(2) According to the processed data set U, the discernibility matrix is generated, and the frequency function value of each attribute C is calculated, and the result is taken as the importance measure of the decision attribute D.

(3) Prepruning of decision the tree: calculate the frequency function value and sum s of each attribute, and get the attribute a with the largest value and its corresponding value r, if r/s is greater than a certain threshold, it reaches the leaf node. Otherwise, a is taken as the current test attribute.

(4) Based on a, a discrete decision tree is constructed, $Root \leftarrow r$ (Root is the root node), then test each possible value v_i of r. Add a new branch corresponding to test $r.\text{value} = v_i$ under $Root \longleftarrow r$. X_{v_i} is the subtable of U that satisfies the attribute value of r as v_i. For each subtable X_{v_i}, if the leaf node is not reached, $AFC4.5\,(X_{v_i}, W - \{a\})$ is called recursively.

The value of each matrix element in the judgment matrix (JM), i.e., the importance of different Table 1 attributes of the same NPE variable, can be determined according to the following table:

According to the knowledge of linear algebra, the maximum eigenvalue of JM $R = (r_{mn})_{ij}$ and its corresponding eigenvector are calculated as follows:

$$R = \lambda_{\max} \times w *. \quad (21)$$

In the above formula (21), $\lambda_{\max}$ is the maximum eigenvalue of JM $R = (r_{mn})_{ij}$, and its corresponding eigenvector $w* = (w_1 *, w_2 *, \cdots, w_n *)$. When experts compare the properties of NPE in two, it is impossible to achieve the same measurement, and there will be some errors. Therefore, to improve the reliability of determining the weight value, it is necessary to check the consistency of the JM.

When the JM $R = (r_{mn})_{ij}$ is completely consistent, $\lambda_{\max} = n$. However, in general, it is difficult to achieve. To test the consistency of the JM, the following formula is needed to calculate its consistency index CI:

$$CI = \frac{|\lambda_{\max} - n|}{n - 1}. \quad (22)$$

In the above formula (22), when $CI = 0$, the JM $R = (r_{mn})_{ij}$ has complete consistency. On the contrary, the larger CI is, the less consistent the JM $R = (r_{mn})_{ij}$ is. To test whether JM $R = (r_{mn})_{ij}$ has satisfactory consistency, it is necessary to compare CI with an average random consistency index RI to get CR, that is,

$$CR = \frac{CI}{RI}. \quad (23)$$

The average random consistency index RI is shown in Table 2.

When $CR < 0.1$, the JM $R = (r_{mn})_{ij}$ has satisfactory consistency; when $CR \geq 0.1$, adjust the JM $R = (r_{mn})_{ij}$ until it is satisfied [19]. At this time, the attribute weights of the variables of the NPE are normalized to get the attribute weights of the variables of the NPE. After the attribute weights of the variables of the NPE are determined, the linear algebraic discretization of the variables of the equation is completed.

2.4. Realize Linear Algebraic Discretization of Variables.

The linear algebraic discretization of a multivariable nonlinear parabolic method is to discretize the whole variable data set. Its basic idea is to fully consider the overall distribution of variables in the attribute space composed of all variable attributes and to use the complementarity and correlation of different attributes in distinguishing objects to generalize the discretization of the variable attribute space of the equation.

Multivariable linear algebraic discretization of NPEs is used to obtain the CBP set. The CBP set is assumed to be empty, and the CBP set for each continuous attribute is added to the CBP set. CBP sets for multivariable linear algebraic discretization include CBPs for all continuous attributes in the data set. After the initial breakpoint set is obtained, the optimal breakpoint is found. The optimal breakpoint is looked up using the objective function. The breakpoint selection of multivariate linear algebraic discretization is the CBP of all variable attributes, considering the complementarity and correlation of variable attributes. After deleting or adding a breakpoint to get the optimal breakpoint, the optimal breakpoint is put into the optimal breakpoint set, and the optimal breakpoint set is initially empty, and the breakpoint is deleted from the initial breakpoint set. In this case, the optimal breakpoint is the best breakpoint for partitioning continuous attributes, and the final breakpoint required for a dataset is in the optimal breakpoint set. In multivariable discretization, the splitting method first finds the optimal breakpoint from all the continuous attributes and then splits the objects in the dataset.

The first step is to find the initial breakpoint. The second step is to find the best breakpoint, and then to divide the continuous attribute values according to the breakpoint.

According to the results of attribute splitting, the variables are merged. In this paper, the method of clustering is used to merge the NPEs. The central idea of grid-based

TABLE 1: Value scale of r_{ij}.

r_{ij}	Definition	Explanation
1	Equal importance	Bothe the indicators of same importance
3	Slightly more important	One of the indicators is little more important
5	Obviously important	One of the indicators is clearly important
7	Much more important	One of the indicators is much more important
9	Extremely important	One of the indicators is extremely important
2468	Between the adjacent judgment	The discount degree of the above two judgments
1/2, 1/4, 1/6, 1/8	Inverse comparison	Inversely compare the two indicators

TABLE 2: The mean random consistency index.

Order number	1	2	3	4	5	6	7	8	9	10	11	12
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49	1.52	1.54

clustering is to give a large set of multivariate data points, which are generally unevenly distributed in the data space. Multivariable linear algebraic discretization is equivalent to the process of hypercube partitioning of variable feature space by hyperplanes perpendicular to different continuous attribute axes. The discrete partition points on each continuous attribute axis correspond to the intersection points of the attribute axis and the hyperplanes divided perpendicular to it. Multivariate linear algebraic discretization is a process in which hyperplanes are determined independently on each continuous attribute axis. The number of hyperplane partitions and the importance of variable attributes in NPE obey certain probability distributions. According to the mathematical probability principle, we can determine the number of hyperplanes in a linear algebraic discretization of different NPEs and use hyperplanes to discretize the variables in the attribute space. Thus, the multivariable linear algebraic discrete method for NPEs is studied.

3. Multivariable Linear Algebraic Discrete Method and its Convergence Analysis

3.1. Multivariable Linear Algebraic Discrete Method. The two new two-layer mesh methods and their convergence analysis for the extended hybrid finite element method for nonlinear parabolic discretization problems. By using the idea of correcting on CG, some new two-layer grid methods can be obtained. From the analysis of convergence of the discrete method, it is found that this method is obviously more effective than the existing two-layer grid method. Applying the idea of a Newton iteration on fine meshes and correction on coarse meshes to the extended hybrid finite element method of nonlinear reaction diffusion problems, a two-layer mesh algorithm is constructed.

The solution of a system of nonlinear equations in a fine space is decomposed into a system of nonlinear equations in

a rough space, and then a Newtonian iterative system of linear equations in a fine space, and then a system of linear equations in a CG as a correction. This method can be divided into the following three steps:

Step 1: On the CG F_H, for any $w_H \in W_H$, $v_H \in \widetilde{V}_H$, $v_H \in V_H$, $(p_H^n - t_H^n, w_H^n) \in W_H \times V_H \times \widetilde{V}_H$ is calculated to satisfy the following nonlinear equations:

$$
\left(\frac{(p_H^n - p_H^{n-1}, w_H)}{\Delta t}, w_H \right) + (\nabla \times \varphi_H^n, w_H)
$$

$$
= (f(p_H^n), w_H)(t_H^n, v_H) \tag{24}
$$

$$
= (p_H^n, \nabla \times v_H)(\varphi_H^n, v_H)
$$

$$
= (K(p_H^n) t_H^n, v_H).
$$

Step 2: On the CG F_h, for any $w_h \in W_h$, $v_h \in \widetilde{V_h}$, $v_h \in V_h$, $(\widetilde{p}n/h - \widetilde{t}n/h, \widetilde{w}n/h) \in W_h \times V_h \times \widetilde{V}_h$ is calculated to satisfy the following nonlinear equations:

$$
\left(\frac{(\widetilde{p}_H^n - \widetilde{p}_H^{n-1}, w_H)}{\Delta t}, w_h \right) + (\nabla \times \widetilde{\varphi}_H^n, w_H)
$$

$$
= (f(p_H^n) + f_p(p_H^n), w_h)(\widetilde{t}_h^n, v_h) \tag{25}
$$

$$
= (\widetilde{p}_h^n, \nabla \times v_h)(\widetilde{\varphi}_h^n, v_h)
$$

$$
= (K(p_h^n) t_h^n, v_h) + (K(p_h^n) t_h^n, v_h).
$$

Step 3: On the CG F_H, for any $w_H \in W_H$, $v_H \in \widetilde{V}_H$, $v_H \in V_H$, $(\overline{p}_H^n - \overline{t}_H^n, \overline{w}_H^n) \in W_H \times V_H \times \widetilde{V}_H$ is calculated to satisfy the following nonlinear equations:

$$
\left(\frac{\overline{p}_H^n - \overline{p}_H^{n-1}, \overline{w}_H}{\Delta t}, w_H \right) + (\nabla \times \overline{\varphi}_H^n, w_H) = (f(\overline{p}_H^n), w_H)(\overline{t}_H^n, v_H) = (\overline{p}_H^n, \nabla \times v_H)(\overline{\varphi}_H^n, v_H) = (T_1, v_H). \tag{26}
$$

Among them,

$$T_1 = K(p_H^n)\bar{t}_H^n + K_p(p_H^n)\bar{p}_H^n t_H^n + K_p(p_H^n)(\bar{p}_h^n - p_H^n)(\bar{t}_h^n - t_H^n).$$

$$(27)$$

Several lemmas need to be quoted again to complete the above proof.

3.2. Lemmas. Symmetric lemma, the integer s is the central axis of symmetry of $u(!s)/i$, and the number of $u(!s)/i$ is equal in the range of equal distance on both sides of the axis of symmetry.

Corresponding lemma if s, t, m, $k$are integers and $s - t = m$, then the number of $u(!t)/i$ in the interval $(t, t + k)$ is equal to the number of $u(!t)/i$ in the interval $(s, s + k)$, that is, the number of two corresponding combinations is equal in the corresponding interval of equal length [1].

Lemma 1. $\lim p_i/\omega_{i,v}^3 = \infty,$ $\quad \lim_{i \to \infty} p_i/\omega_{i,v}^2 = \infty \lim_{i \to \infty}$ $p_i/\omega_{i,v} = \infty.$
Prove:

(1) $\because \omega_{i,v} = A_i/B_i$
Set up $a_i = p_i/\omega_{i,v}^3 = (B_i/A_i)^3 p_i$, $a_{i+1} = p_{i+1}/\omega_{i+1,v}^3 = (B_{i+1}/A_{i+1})^3 p_{i+1}$
$a_{i+1} - a_i = p_{i+1}(B_{i+1}/A_{i+1})^3 - p_i(B_i/A_i)^3 = p_{i+1}(B_i/A_i)^3$
$(p_{i+1} - 1/p_{i+1})^3 - p_i(B_i/A_i)^3$
$= (B_i/A_i)^3(p_{i+1} \cdot (p_{i+1} - 1/p_{i+1})^3 - p_i)$
$(p_{i+1} - 1)^3/p_{i+1}^2 = p_{i+1}^3 - 3p_{i+1}^2 + 3p_{i+1} - 1/p_{i+1}^2 = p_{i+1} - 3 + 3/p_{i+1}$
$\therefore p_{i+1} = p_i + \omega_{i,v}, \omega_{i,v} > 3$

$$\therefore p_{i+1} \cdot \left(\frac{p_{i+1} - 1}{p_{i+1}}\right)^3 - p_i = p_{i+1} - 3 + \frac{2}{p_{i+1}} - p_i$$

$$= \omega_{i,v} - 3 + \frac{2}{p_{i+1}} > 0.$$

$$(28)$$

(2) $\therefore (B_i/A_i)^3 > 0, \therefore a$
$_{i+1} - a_i > 0, \therefore a_{i+1} > a_i \therefore p_{i+1}(B_{i+1}/A_{i+1})^3 > p_i(B_i/A_i)^3.$

(3) $\therefore \lim_{i \to \infty} p_i/\omega_{i,v}^3 = \infty, \because \omega_{i,v}^3 > \omega_{i,v}^2 > \omega_{i,v}, \therefore \lim_{i \to \infty} p_i/\omega_{i,v}^2 = \infty \lim_{i \to \infty} p_i/\omega_{i,v} = \infty.$

Lemma 2. *The number* $\delta_{\gamma,\beta}^{(!s!-t)}$ *of combination* $u_\gamma^{(!s!-t)}$, *if* $p_\gamma^2 < \beta < p_{\gamma+1}^2$, $\omega_{n,v}p_n = \beta,$ *then in the* ∞ *range,* $\delta_{\gamma,\beta}^{(!s!-t)} = p_n/\omega_{n,v}.$
Prove:

(1) *The number* $u(!s)/i$ *of combination* $\delta_{i,\beta}^{(!s)}$.
For the convenience of narration, a is the number of $u(!s)/i$ *in the interval* $(0, s)$, *B is the number of* $u(!s)/i$

in the interval (s, β), *c is the number of* $u_i^{(!0)}$ *in the interval* $(-s, 0)$, *and D is the number of* $u_i^{(!0)}$ *in the interval* $(0, \beta - s)$.

$\because s < p_\gamma$, *According to closed lemma 2, E = 1.*

According to the principle of symmetry, C = E = 1,
$\therefore D + C = D + E = D + 1$.

According to the corresponding principle, A = C, B = D is obtained.

$$\therefore A + B = C + D = D + 1 - j. \qquad (29)$$

According to the above definition.
$\delta_{\gamma,\beta-s}^{(!0)} = D$, $\delta_{\gamma,\beta}^{(!S)} = A + B = C + D = D + 1 = \delta_{\gamma,\beta-S}^{(!S)} + 1$

(2) *Number* $\delta_{\gamma,\beta}^{(!-t)}$ *of combination* $u_\gamma^{(!-t)}$.
Similarly.

$\because t < p_\gamma$, *According to closed lemma 2, C = 1.*

According to the correspondence principle, B = D

$$\delta_{\gamma,\beta+T}^{(!0)} = D + C. \qquad (30)$$

(3) *The number of nontwo kinds of remainder combination is infinite.*

$\because S$ *and T are constants and* β *is variable,*

$$\lim_{\beta \to \infty}(\beta - s) = \beta \lim_{\beta \to \infty}(\beta + t) = \beta. \qquad (31)$$

$\therefore$*In the range of infinity,*

$$\delta_{\gamma,\beta}^{(!s)} = \delta_{\gamma,\beta}^{(!0)}\delta_{\gamma,\beta}^{(!-t)} = \delta_{\gamma,\beta}^{(!0)}. \qquad (32)$$

When,

$$\omega_{n,v}p_n = \beta\delta_{\gamma,\beta}^{(!0)} = p_n\delta_{\gamma,\beta}^{(!s!-t)} = \frac{p_n}{\beta} \cdot \frac{p_n}{\beta} \cdot \beta = \frac{p_n}{\omega_{n,v}}. \qquad (33)$$

$\because \lim_{\gamma \to \infty} p_n/\omega_{n,v} = \infty, \therefore \lim_{\gamma \to \infty} \delta_{\gamma,\beta}^{(!s!-t)} = \infty.$

That is to say, in the range of ∞, the number of all nontwo types of remainder combinations is infinite.

The proof of lemma is finished.

3.3. Convergence of Mesh Discretization Method. Different results can be obtained under the grid settings, with large differences. If these results are in good agreement, it indicates that the simulation results are stable and reliable. On this basis, the convergence of the three-step two-layer grid discretization method is analyzed.

Order
$\hat{p}n/h = \tilde{p}n/h + \bar{p}n/H, \hat{t}n/h = \tilde{t}n/h + \bar{t}n/H, \hat{\varphi}n/h = \tilde{\varphi}n/h + \bar{\varphi}n/H$

It can get the following equation:

$$\left(\frac{a^n - a^{n-1}}{\Delta t}, w_h\right) + \left(\nabla \times \left(\prod_h \varphi^n - \varphi_h^n\right), w_h\right) = (E + G, w_h)(\beta^n, v_h) = (a^n, \nabla \times v_h)(\gamma^n, v_h) = \left(\prod_h \varphi^n - \varphi^n, v_h\right) + (K(p^n)t^n, v^h) - (T_4, v_h). \quad (34)$$

From the above theorem analysis and convergence analysis, it can be seen that, as long as the step H of the coarse space selected in the above algorithm satisfies $H = O(h_{3k+1}^{k+1})$, the two-layer mesh method established can maintain the optimal approximation of the solution of the mixed finite element method. The idea is applied to the reaction diffusion problem when the reaction term is a pressure p-related term ($f(p, \nabla p)$), and the tensor K is a nonlinear term related to p. That is $K(p)$. That is, to solve the original complex problem on the coarse mesh, and then to carry out Newton iteration on the fine mesh, that is to say, to solve a relatively simple problem on the fine mesh using the extended mixed finite element method, and then to correct it on the coarse mesh. By convergence analysis, the method can keep the best approximation of the mixed finite element solution.

4. Conclusions

NPE has important use value in many fields, such as science and technology manufacturing. But NPE usually contains many variables and is difficult to solve. NPE is an important branch of mathematics. It is one of the important tools needed to solve the linear algebraic discrete method of multivariable in reality. Based on the existing research results for NPE, the multivariable linear algebraic discrete method for NPE is studied. First, the NPE is discretized, the algebraic form of the nonlinear equation is transformed into the vector form, and the RS and IE are constructed to allocate the weights of different variable attributes. According to the given variable attribute weight, the multiple variables in the equation are discretized by linear algebra. Through the above steps, the research on the multivariable linear algebraic discretization method of NPE is completed in order to provide some help for the research in this field.

Data Availability

The data used to support the findings of this study are available from the author upon request (zuoli@mjc-edu cn).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] A. K. H. Sedeeg, R. I. Nuruddeen, and J. F. Gómez-Aguilar, "Generalized optical soliton solutions to the (3+1)-dimensional resonant nonlinear Schrödinger equation with Kerr and parabolic law nonlinearities," *Optical and Quantum Electronics*, vol. 51, no. 6, pp. 173–215, 2019.

[2] M. Kaur, S. Kadam, and N. Hannoon, "Multi-level parallel scheduling of dependent-tasks using graph-partitioning and hybrid approaches over edge-cloud," *Soft Computing*, vol. 26, no. 11, pp. 5347–5362, 2022.

[3] G. Ricarte, R. Teymurazyan, and J. M. Urbano, "Singularly perturbed fully nonlinear parabolic problems and their asymptotic free boundaries," *Revista Matemática Iberoamericana*, vol. 35, no. 5, pp. 1535–1558, 2019.

[4] J. Z. Cheng, L. Y. Jin, and M. FangS, "Properties of solutions for a class of nonlinear pseudo parabolic equations," *Advances in Applied Mathematics*, vol. 09, no. 04, pp. 551–559, 2020.

[5] C. Eichenberg, "Special solutions to a nonlinear coarsening model with local interactions," *Journal of Nonlinear Science*, vol. 29, no. 4, pp. 1343–1378, 2018.

[6] T. Nishino and T. Yokota, "Effect of nonlinear diffusion on a lower bound for the blow-up time in a fully parabolic chemotaxis system," *Journal of Mathematical Analysis and Applications*, vol. 479, no. 1, pp. 1078–1098, 2019.

[7] A. J. Castro, S. Rodriguez-Lopez, and W. Staubach, "L²-solvability of the Dirichlet, Neumann and regularity problems for parabolic equations with time-independent Hölder-continuous coefficients," *Transactions of the American Mathematical Society*, vol. 370, no. 1, pp. 265–319, 2017.

[8] M. Kumar and A. K. Tiwari, "Some group-invariant solutions of potential Kadomtsev-Petviashvili equation by using Lie symmetry approach," *Nonlinear Dynamics*, vol. 92, no. 2, pp. 781–792, 2018.

[9] Z. Wang, "Inverse source problem of time dependent nonlinear parabolic equation," *Journal of Lanzhou University of Arts and Science(Natural Sciences)*, vol. 36, no. 4, pp. 1–9, 2022.

[10] H. Liu, R. P. Zhang, and J. R. Huo, "Finite difference method based on fast sine discrete transformation for solving semi linear parabolic equation," *Advances in Applied Mathematics*, vol. 9, no. 12, pp. 2209–2216, 2020.

[11] M. Kaur and S. S. Kadam, "Discovery of resources using MADM approaches for parallel and distributed computing," *Engineering Science and Technology, an International Journal*, vol. 20, no. 3, pp. 1013–1024, 2017.

[12] S. S. Kevorkyants, "One solution of the forward problem of dc resistivity well logging by the method of volume integral equations with allowance for induced polarization," *Izvestiya - Physics of the Solid Earth*, vol. 54, no. 2, pp. 359–371, 2018.

[13] T. Q. A. C. Zhang and C. J. Zhang, "A general class of one-step approximation for index-1 stochastic delay-differential-algebraic equations," *Journal of Computational Mathematics*, vol. 37, no. 2, pp. 151–169, 2019.

[14] B. P. Ouyang and Y. F. Li, "Study on blow up of solutions of a class of Parabolic Equations with nonlinear gradient term,"

Journal of Chongqing Normal University(Natural Science), vol. 37, no. 5, pp. 83–88, 2020.

[15] B. Martens and M. Gerdts, "Convergence analysis of the implicit euler-discretization and sufficient conditions for optimal control problems subject to index-one differential-algebraic equations," *Set-Valued and Variational Analysis*, vol. 27, no. 2, pp. 405–431, 2019.

[16] A. Jadhav, M. Kaur, and F. Akter, "Evolution of software development effort and cost estimation techniques: five decades study using automated text mining approach," *Mathematical Problems in Engineering*, Article ID 5782587, pp. 1-17, 2022.

[17] L. M. Anguas, M. I. Bueno, and F. M. Dopico, "A comparison of eigenvalue condition numbers for matrix polynomials," *Linear Algebra and Its Applications*, vol. 564, no. 2, pp. 170–200, 2019.

[18] M. Kaur, A. Jadhav, and F. Akter, "Resource Selection from Edge-Cloud for IIoT and Blockchain-Based Applications in Industry 4.0/5.0," *Security and Communication Networks*, 2022.

[19] H. Q. Lu and Z. C. Zhang, "Blowup time estimates for a parabolic p-Laplacian equation with nonlinear gradient terms," *Zeitschrift für Angewandte Mathematik und Physik*, vol. 70, no. 3, pp. 90–132, 2019.

Research Article

Antioxidant Effect and Acute Oral Toxicity of Hot Springs

Israa J. Hakeem [ID][1] and Gashaw Tadele Zewudie [ID][2]

[1]*University of Jeddah, College of Science, Department of Biochemistry, Jeddah, Saudi Arabia*
[2]*Department of Chemistry, Mizan-Tepi University, Ethiopia*

Correspondence should be addressed to Gashaw Tadele Zewudie; gashaw@mtu.edu.et

According to the research, there are many illnesses for which therapeutic mineral hot springs are employed as an alternative. Its physicochemical characteristics have a substantial body of evidence. The in vivo antioxidant effect of Mosul's hot springs in Iraq has been investigated in the current investigation. An experimental design for toxicity, a control group, and a study group were created. In addition, *in vivo* antioxidant effect of the hot springs of Mosul, Iraq, has been studied by the lipid antiperoxidation method with ($p < 0.05$), in vitro by the free radical scavenging method (DPPH) for its complexing capacity of hot springs. In acute oral toxicity in vivo at fixed doses, looking for signs and symptoms of toxicity, there are no signs of intoxication or significant changes in the biochemical analysis (blood count). And, it was discovered that the variances are substantial. The animal was necropsied, and hematological and biochemical parameters were determined, as well as the organs' histological processing at the study's conclusion. It was found that the thermal waters from Mosul, Iraq, are medicinal mineral waters, chlorinated, sodium, and sulfated, nontoxic and have an antioxidant effect. With the help of the research's findings, it is hoped to provide scientific support for knowledge that, when made public, encourages the development of Mosul's hot springs as a safe and environmentally friendly tourist destination. With the results of this research, the parameters were presented with their mean and standard deviation statistics, promoting the ecological and sanitary tourism development of the Mosul hot springs, which will generate more significant income for the population, therefore growth in the region.

1. Introduction

Since mineral waters have been a part of our existence since ancient times, understanding its origin and mineralization via an investigation of the formation mechanisms that give birth to this water's source and their interaction with the formations in which it is collected has tremendous value [1]. Such is the case that it was established in 2013 [2, 3], as a natural resource, for the treatment of various diseases. Groundwater and its hydrogeological context are related to physical-chemical and biological geological factors. [4, 5] Within the hydrological cycle, waters with characteristics can be differentiated by the degree of mineralization and/or by their temperature from groundwater, called mineral waters [6]. The intake of natural mineral water containing sulfates leads to gallbladder contraction, probably induced by the release of cholecystokinin [7]. Water toxicity can be determined by applying the 16-day repeated dose oral toxicity test in rodents; this method provides information on the health risk likely to arise from exposure to the test substance by oral administration [8]. This guide is primarily intended for use with rodents, preferably rats.

Some diseases related to free radicals are increased in our environment because the body's antioxidant defences are overcome by oxidative attack due to the work environment or other conditions that give rise to oxidative stress [9]. These disorders are caused by the oxygen's damaging effects, especially those of its free radicals in the mitochondria, which target all the elemental constituents and create various metabolites like malondialdehyde [10]. There are a variety of cases of acute intoxication, which manifested with a pathological clinical picture, after a single exposure to a substance or multiple exposures in a period of 24 hours [11]. For all these reasons, it was decided to study the antioxidant

effect and acute toxicity of the hot springs of Mosul, Iraq, determining the antioxidant effect according to the lipid antiperoxidation method, the acute oral toxicity, the concentrations of sodium, potassium, calcium, magnesium, sulfates, chlorides, and other corresponding in the four seasons of the year, to provide scientific support for its use as mineral-medicinal thermal water.

2. Materials and Methods

The water sample was collected from the thermal waters from Mosul hot spring at an altitude of 36.361312′, 43.120445′ longitudes. To preserve and avoid decomposition, it was transported to the laboratory at a temperature of 4–8°C. To carry out the analyses, the techniques of the standard methods for the examination of water and wastewater are a joint publication of the American Public Health Association (APHA), 23rd edition [12]. The peristaltic pump is used to convey the liquid sample to the nebulizer system, where it is converted into an aerosol by the action of argon gas before being sent to the ionization zone for ICP-MS analysis [13]. A high-frequency current induces an oscillating magnetic field that creates a plasma from a flow of argon gas. Plasma can reach temperatures as high as 8000 K. The atoms in the sample are ionized under these circumstances. An increasing vacuum interface separates the charged and noncharged particles in the quadrupole filter. The sample's abundance of each tuned mass is determined at the detector. In ion exchange chromatography, the charge characteristics of molecules are used to separate ions and polar molecules for further analysis. Large proteins, tiny nucleotides, and amino acids are all examples of charged molecules that can benefit from this technique. The injectable solution is referred to as a "sample," and the constituents are referred to as "analytes." Anions are the analytes in this scenario (chlorides, sulfates).

2.1. Determination of Toxicity at Repeated Doses for 16 days by Oral Administration in Rats.
20 Wistar rats were used in healthy young adult nulliparous and nonpregnant females commonly used laboratory strains Group 1 control (autoclaved drinking water) consisting of five females and five males and Group 2 study (thermal water of Mosul, Iraq) consisting of five females and five males. The trial was conducted for 16 days and administered for four days. A dose was used for the test sample, 1000 mg/kg body weight. The volume to be administered every four days was adjusted, which depended on the study group's body weight variations. In addition, the limit volume (of aqueous solutions) for this species of 2 ml/100 g of body weight was considered.

Maintenance and Feeding: the experimental animals were adapted to the bacterium's conditions for four days before starting the test; it is distributed individually in polycarbonate boxes type $T3$ (plastic) with grid-bottom metal.

Water and food were sterilized with an autoclave to avoid any contamination. The environment's temperature was conditioned to $23 \pm 2.1°C$, relative humidity <70%, and photoperiod of 12 hours of light and 12 hours of darkness.

The clinical and pattern signs of behaviors related to the toxicity of the experimental animals were observed daily, as well as changes in the skin, fur, mucous membranes, eyes, hyperventilation, vasodilation, restlessness, and somatic motor activity. Special attention was paid to the possible occurrence of some crucial signs for the study, such as tremor, convulsion, diarrhea, lethargy, irritability, salivation, sleep, and coma.

Daily and four days measurements of water and food consumption were made. Individual body weights were determined the day before the hot springs and drinking water administration, then at the 3, 8, 10, and 15 days that the rehearsal lasts. Hematological and biochemical parameters (blood count, urea, creatinine, TGO, TGP) were determined 16 days after being treated with the sample problem; blood samples were extracted by intracardiac puncture from the previously anesthetized and fasting animals.

On the last day of the experimentation, the animals of the study group and the control group were sacrificed for cervical dislocation after anesthesia. Then, the autopsy was performed where the body surface, cavities, and organs were examined.

2.1.1. Evaluation of the Antioxidant Effect

(1) Technique for the Detection of Lipid Peroxidation Products (TBARS). For this process, the livers of female albino rats were used, weighing 200–220 g, and kept in an acrylic cage, to which balanced food and water were provided ad libitum, stopping feeding them 12 hours before the experiment. The animals were divided into two groups [14]. The five rats in the control group were given 5 mL/kg of body weight of water suitable for human consumption, while 15 rats, divided into groups of 5 rats, were given hot springs at a rate of 3, 5, and 9 mL/kg of weight, respectively. Then, 1 g of the liver was separated and homogenized in 10 volumes of 0.154 M KCl, and subsequently, the production of the malondialdehyde-thiobarbituric acid complex (MDA-TBA) was measured at $535 \mu m$ in a UV–visible spectrophotometer [15].

During the first 30 minutes of the first four hours after injection and then every day for the next 16 days of the experimental study, individual animals were watched for any indications or symptoms of toxicity. The evaluation's goal was to determine the cause of death and the exact moment that poisoning signs and symptoms, weighing 28 to 32 g. They were stopped feeding 12 h before the experiment.

Five mice for the control group were dosed with a dose of 20 mL/kg body weight of 0.9% sodium chloride. The study group consisted of 5 groups of 5 mice each. As we worked with five different doses: 10, 20, 30, 40, and 50 mL/kg, each group received only one of them, respectively. (7). The animals were observed individually during the first 30 minutes of the 4 hours following administration and then daily until the 16 days of the experimental trial, looking for

signs and symptoms of toxicity. The evaluation was aimed at determining death and time of occurrence of signs and symptoms of toxicity, including its onset and duration, as well as changes in the skin, mucous membranes, and eyes, in the respiratory, circulatory, central nervous, and autonomic systems, in somatomotor effect and behavior. Special attention was paid to the potential occurrence of seizures, salivation, diarrhea, lethargy, drowsiness, and coma. At the end of the period, for assessing any toxicity-related symptoms and indicators, the animals [16] were sacrificed and subjected to necropsy in which the macroscopic pathological changes of organs and tissues were observed, mainly in the heart, lung, liver, spleen, and kidney. Subsequently, a histopathological examination of the liver and kidney was performed.

The description of variables is expressed in means and standard deviation; statistical comparison of groups, using one-way ANOVA followed by Tukey's post hoc test, was considered significant with $p < 0.05$ at the 95% confidence interval. The statistical program SPSS version 17 was used.

3. Results

Table 1 shows the results of the analytical determinations of total metals (cations), during the four seasons of the year 2022, of the thermal water of Mosul. Table 2 and Figure 1 show the descriptive statistics of the metal results, and Tables 3 and 4 show the average values of cations and anions in (mg/L), meq/L, and % meq.

Table 5 describes the results of the observation of the experimental animals for 16 days of testing for antioxidant study.

3.1. Determination of Toxicity at Repeated Doses for 16 days by Oral Administration in Rats

3.1.1. Clinical Observations. The observations of the clinical signs of the animals were made daily; fundamentally, the physical state, behavior in the nasal and ocular mucous membranes, secretions, or alterations related to toxicity were sought [17]. Palpation of the abdomen was performed without finding changes, and special attention was paid to the possible occurrence of signs such as tremors, convulsions, diarrhea, lethargy, salivation, sleep, and coma. And at the end of the experimentation, no changes or alterations in the clinical signs were observed. They considered the expected behavior of the animals of said species. The animals were weighed at 1, 4, 8, 12, and 16 days. The results show that the tendency to increase body weight was constant during the study in both groups (Figure 2).

In Figure 3, it can be seen that when comparing the study group and the control group with time, there is no significant difference; no effect of the groups and time on the weights of Wistar rats was found.

3.2. Clinical Laboratory Tests. Hematological and blood chemistry determinations were carried out 16 days after gastric gold administration of the hot springs. Blood collection was performed by cardiac puncture after fasting for 12 h of the animals; safety vials were collected. In both cases, the parameters were presented with their mean and standard deviation statistics, which groups established. With the statistical treatment, no changes or alterations in blood formula were evident. In the anatomopathological study, at the end of the treatment, there was no evidence of the death of any animal, so euthanasia was performed on the 20 experimental animals in which ether was previously applied [18]. Then, cervical dislocation was performed in order to provide the necessary macroscopic observations for later microscopic research and then proceeded to the extraction and weighing of the following organs: heart, lungs, kidneys, spleen, and liver. No significant difference was observed concerning the weights of the organs; in addition, in the microscopic histological study that was carried out on the lung, heart, spleen, and liver, no macroscopic or microscopic alterations were attributable to the sample were found under study. There is no significant difference in the weights of the organs of the control group concerning the study group. There was no animal death during the study, so all were sacrificed at the end of the investigation.

4. Discussion

The studies carried out for the chemical analysis of the hot springs of Mosul, Iraq, reported the presence of 0.2871 g sodium sulfate, 3.1955 g sodium chloride, etc. In 2018, the Governorate of Mosul requested the Water Research Laboratory for the chemical analysis of four water samples [19]. The test report reports that the following potassium was 57.69 mg/L, and sodium was 1350 mg/L. Al-Shahri [20] states that "the thermal and mineral waters of Mosul are sodium chloride and sulfated." Hashim, in 2020, [21] in his study of the chemical and toxicological evaluation of lithium, found 15.95 mg/L of said element in the thermal water. Hashim [21] affirms that the thermal water of Mosul presents sodium 1682.31 mg/L, chloride 1962.31 mg/L, and sulfate 861.91 mg/L; however, in the study that was carried out, sodium 1869.63 mg/L, sulfate 7796.30 mg/L, potassium 75.06 mg/L, chloride 1610.62 mg/L, and by concerning previous studies, sodium concentrations increased by 27% and potassium by 24% over time. However, the concentrations of sulfates decreased by 11% and chlorides by 20%, decreasing as time went by, probably due to the degradation of volcanic rocks, among others. Table 1 shows the results of some total metals above the detection limits (DL) of the 52 metals analyzed during the four seasons of the year 2022, highlighting some metals of toxicological interest: arsenic, cadmium, mercury, lead, sodium, and potassium but when making comparisons of the values in different seasons of the year, there is no significant difference between them.

Table 2 shows that the mean of the metals in the summer season is higher. The mean values of the metals corresponding to the spring are lower and probably vary with the telluric movements, degradation of the internal rocks, or, failing that, with the rains that drag minerals to the water table; however, these differences are not significant

TABLE 1: Total metals of the thermal water of Mosul during the four seasons.

Total metals	DL	M1 Summer	M2 Fall	M3 Spring	M4 Winter
Aluminium (mg/L)	0.0011	0.0396	BDL	BDL	BDL
Antimony (mg/L)	0.006237	0.00704	0.006237	0.005687	0.005775
Arsenic (mg/L)	0.000033	0.053108	0.056221	0.052426	0.054472
Barium (mg/L)	0.00011	0.04202	0.04213	0.04191	0.04191
Boron (mg/L)	0.0022	50.9509	50.1479	51.7979	51.2402
Cadmium (mg/L)	0.000011	BDL	BDL	BDL	BDL
Calcium (mg/L)	0.0033	254.9041	247.8619	244.233	243.1836
Caesium (mg/L)	0.00011	2.83426	2.92952	2.87221	2.95152
Cobalt (mg/L)	0.000011	0.000429	0.000418	0.000451	#Value!
Strontium (mg/L)	0.00022	6.73464	6.34711	6.41311	6.29211
Germanium (mg/L)	0.00022	0.01221	0.01188	0.01243	0.01177
Lithium (mg/L)	0.00011	24.55926	24.28162	24.09143	22.99561
Magnesium (mg/L)	0.0011	57.8523	59.0073	58.2835	59.2636
Manganese (mg/L)	0.000033	0.224235	0.207306	0.202367	0.101167
Mercury (mg/L)	1.1	BDL	BDL	BDL	BDL
Molybdenum (mg/L)	0.000022	0.004972	0.004422	0.004411	0.004323
Lead (mg/L)	0.00022	BDL	BDL	BDL	BDL
Potassium (mg/L)	0.044	86.02	83.149	78.562	81.939
Rubidium (mg/L)	0.00033	0.30635	0.28732	0.28677	0.28204
Selenium (mg/L)	0.00044	0.00275	0.00242	0.00319	0.00231
Silica (mg/L)	0.099	49.203	49.302	49.61	51.205
Silica (mg/L)	0.044	23.001	23.045	23.188	23.936
Sodium (mg/L)	0.0066	1,803.19	1,72,666	1,762.05	1,769.70
Thallium (mg/L)	0.000022	0.000363	0.000352	0.000352	0.000286
Titanium (mg/L)	0.00022	0.02618	0.02497	0.02387	0.0242

*BDL: below detectable limit.

TABLE 2: Descriptive data of the total metals found in Mosul's thermal water throughout the course of the four seasons.

Station	Statistical	P
Half	85.82	
Summer medium	0.0483	
Minimum	0	
Maximum	1639.26	
Half	82.67	0.99*
Medium spring	0.0511	
Minimum	0	
Maximum	1569.69	
Half	83.7	
Medium autumn	0.0477	
Minimum	0	
Maximum	1601.86	
Half	84.12	
Medium winter	0.0495	
Minimum	0	
Maximum	1608.82	

FIGURE 1: Annual behavior of metal concentrations.

TABLE 3: The four stations' average cation values.

Cation	mg/L	meq/L (%)
Sodium	1,604.91	69.78 (77.1%)
Potassium	74.925	1.94 (2.14%)
Lithium	21.8	3.15 (3.47%)
Calcium	225.04	11.23 (12.41%)
Magnesium	53.27	4.39 (4.85%)
Total	1,979.94	90.49 (100%)

$p > 0.005$. Figure 1 shows that the averages of the concentrations tend to decrease. However, this difference is not statistically significant or the variation between the year's seasons with $p > 0.005$.

In addition, Table 4 shows the values of the concentrations of the chemical elements that indicate the classification of the waters and their chemical equivalent of biochemical and therapeutic interest, considering that the concentration is more significant than the concentration of 1 g/L or 80 meq/L has therapeutic applicability. When an ion is greater than 20 mEq/L, this gives the water its name. According to the predominant ionic composition, mineral waters are classified as chlorinated, sulfated, and sodium waters.

TABLE 4: The four stations' average anion values.

Anions	mg/L	meq/L (%)
Chloride	1,829.83	51.62 (72%)
Bicarbonate	253.63	4.16 (5.8%)
Nitrate	0.04	0 (0%)
Sulfate	763.97	15.91 (22.18%)
Total	2,847.47	71.69 (100%)

TABLE 5: Result of the observation of experimental animals for 16 days at a dose of 20 mL/kg.

Observation period	Clinical signs
Half an hour after administration	Slight drowsiness and grouping in the center of the cage
From 3 hours after administration	Signs of recovery
From 4 hours after administration	Normal postural reflex, normal grooming habits, and consumption of food and water
Second day	Intermittent bunching at cage ends
From the third day, until the end of the study	No signs of toxicity were observed, and the animals were normal in their behavior and lifestyle
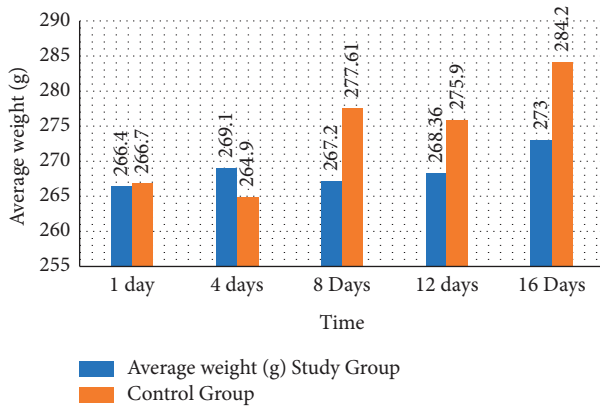
FIGURE 2: Variation of the body weight of the Wistar rats that received the thermal water of Mosul orally for 16 days.
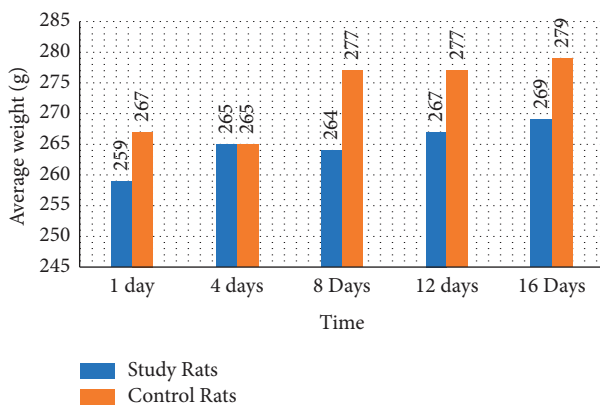
FIGURE 3: Behavior of the mean weight values of Wistar rats from the control and study groups during the 16 days of the study.

In Table 5, it is observed that only half an hour after the administration of the 20 mL/kg dose, the animals presented clinical signs, slight drowsiness, and grouping in the center of the cage. Subsequently, signs of recovery, normal postural reflex, normal grooming habits, and consumption of food and water were observed until day 16. However, in Figure 1 as demonstrated by the multivariate statistical model of the influence of the control group, study, and time on weights over the 16 days of research, this difference is not statistically significant. In addition to the aforementioned statement, it may be inferred that, even if there are no overt clinical changes, it is possible that the thermal waters would be having a purgative impact on rats due to the loss in weight. There was no evidence of a group or time effect on weights. You can see the difference between the weight averages between the first and second dose times, where the study group's values start to slow down weight gain while the control group's values continue to gain weight over the course of the study.

On the other hand, the histological sections of the different organs (liver, kidney, and heart) of the mouse presented alterations typical of foreign agents, which when entering the organism modifies the cellular physiology, thus microvascularization was observed in hepatocytes, altered bile ducts, and glomerulonephritis to a lesser extent, while the heart remained normal. These alterations regenerate, and it demonstrates that at the dose given, they may not typically entail permanent harm, so it would be advisable to carry out new studies that evaluate chronic toxicity. NaCl 0.9% 5 mL/kg at 0.0527TBARS hm, for ATP3 0.0430TBARS hm; for ATP5 0.0418TBARS hm, and for ATP9 0.039TBARS hm observed. The results show that there is a significant difference between the results of the control group (NaCl 0.9% 5 mL/kg of weight) and those of the study groups (hot springs 3, 5, and 9 mL/kg of weight, respectively), showing that there is an inversely proportional relationship between the dose administered and the absorbance reading reached (lipid peroxidation). This demonstrates the antioxidant activity of the studied waters.

In his study, Eixarch, magnesium sulfate can relieve oxidative stress and reduce inflammatory cytokines in the placenta of rats from intrahepatic cholestasis to the pregnancy model. They concluded that $MgSO_4$ had a beneficial effect on improving the growth of offspring in the rat model

of PCI [22]. The protective effect of MgSO$_4$ in relieving oxidative damage and the inflammatory response in the placenta may play an essential role in the process. MgSO$_4$ can improve the placenta's function. In our study, magnesium and sulfate were determined in considerable concentrations that would probably produce a protective effect against oxidative stress. As far as toxicity is concerned, the study was carried out to determine the toxicity of the thermal waters of Mosul, Iraq, administered to Wistar rats by orogastric route once a day for 16 days.

Biochemical parameters such as creatinine are considered markers of kidney damage which are directly related to body mass; it is evident that the rats in the study group (100 mg/kg) decreased creatinine values; these values coincide with the loss of body weight during the second dose of treatment. Agree with the values found in our study that indicate there are no signs of intoxication.

5. Conclusion

The components and chemical concentrations of the thermal water of Mosul, Iraq-Puno, were determined in the year's four seasons, being an average of the highest concentrations, Na$^+$ 1682.31 mg/L, Cl$^-$, 864.32 mg/L, SO$_4^{2-}$ 801.61 mg/L and K$^+$, 80.21 mg/L. The chemical components of the thermal water of Mosul, Iraq, have the highest average concentration in summer (86.31 mg/L) and the lowest average concentration (83.07 mg/L) in spring. The thermal water of Mosul, Iraq, is not toxic in Wistar rats as it is administered orally for 16 days.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

[1] Y. Al-Shahri, B. Maulood, and J. Toma, "A study on a sulfur spring (ain Al kibrit) ecosystem along tigris river Mosul, Iraq," *Journal of Advanced Laboratory Research in Biology*, vol. 7, p. 6, 2016.

[2] K. Al-Youzbakey and A. Sulaiman, *The Evaluation of Chemical Characterization for Selected Wells Water in Mosul – Bahshiqa – Shalalat Area*, Ninivah Governorate, Northern Iraq, 2018.

[3] R. Baird and L. Bridgewater, *Standard Methods for the Examination of Water and Wastewater*, American Public Health Association, Washington, DC, USA, 23rd edition, 2017.

[4] A. Bulgheroni, A. Kinsner-Ovaskainen, S. Hoffmann, T. Hartung, and P. Prieto, "Estimation of acute oral toxicity using the No Observed Adverse Effect Level (NOAEL) from the 28 day repeated dose toxicity studies in rats," *Regulatory Toxicology and Pharmacology*, vol. 53, no. 1, pp. 16–19, 2009.

[5] A. A. Hamad, M. L. Thivagar, J. Alshudukhi et al., "Secure complex systems: a dynamic model in the synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9719413, 6 pages, 2021.

[6] E. Eixarch, F. Figueras, E. Hernández-Andrade et al., "An Experimental model of fetal growth restriction based on selective ligature of uteroplacental vessels in the pregnant rabbit," *Fetal Diagnosis and Therapy*, vol. 26, no. 4, pp. 203–211, 2009.

[7] N. Eizadi-Mood, R. Heshmat, R. Meamar, and N. Motamedi, *The Relative Risk of the Toxico-Clinical Parameters Based on Poisoning Severity and Outcomes in Patients with Acute Poisoning*, 2021.

[8] N. Even-Tzur, Y. Kloog, M. Wolf, and D. Elad, "Mucus secretion and cytoskeletal modifications in cultured nasal epithelial cells exposed to wall shear stresses," *Biophysical Journal*, vol. 95, no. 6, pp. 2998–3008, 2008.

[9] A. S. Al-Obeidi, S. Fawzi Al-Azzawi, A. Abdullah Hamad, M. L. Thivagar, Z. Meraf, and S. Ahmad, "A novel of new 7D hyperchaotic system with self-excited attractors and its hybrid synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 3081345, 11 pages, 2021.

[10] B. Gammelgaard and B. Packert Jensen, "Application of inductively coupled plasma mass spectrometry in drug metabolism studies," *Journal of Analytical Atomic Spectrometry*, vol. 22, no. 3, p. 235, 2007.

[11] M. Alsaffar, A. A. Hamad, A. Alshammari et al., "Network management system for IoT based on dynamic systems," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 9102095, 8 pages, 2021.

[12] N. Abdalwahid Hashim and M. Sabah Kassim, "A comparative study about lithium bromide-water and lithium chloride-water solutions in pumpless absorption solar cooling systems in iraqi circumstances," *Journal of Engineering and Sustainable Development*, vol. 24, no. 4, pp. 79–88, 2020.

[13] S. Z. Jassim, R. Raiswell, and S. H. Bottrell, "Genesis of the Middle Miocene stratabound sulphur deposits of northern Iraq," *Journal of the Geological Society*, vol. 156, no. 1, pp. 25–39, 1999.

[14] L. Lidfors, A. Wichman, B. Ewaldsson, and A.-S. Lindh, "Enriched cages for groups of laboratory male rats and their effects on behaviour, weight gain and adrenal glands," *Laboratory Animals*, vol. 48, 2013.

[15] V. Lobo, A. Patil, A. Phatak, and N. Chandra, "Free radicals, antioxidants and functional foods: impact on human health," *Pharmacognosy Reviews*, vol. 4, no. 8, pp. 118–126, 2010 Jul.

[16] M. C. Albertini, M. Dacha, L. Teodori, M. E. Conti, and M. Enrique, "Drinking mineral waters: biochemical effects and health implications - the state-of-the-art," *International Journal of Environment and Health*, vol. 1, pp. 153–169, 2007.

[17] N. Marquardt, M. Feja, H. Hünigen et al., "Euthanasia of laboratory mice: are isoflurane and sevoflurane real alternatives to carbon dioxide?" *PLoS One*, vol. 13, no. 9, Article ID e0203793, 2018.

[18] E. Nagababu, J. M. Rifkind, S. Boindala, and L. Nakka, "Assessment of antioxidant activity of eugenol in vitro and in vivo," *Methods in Molecular Biology*, vol. 610, pp. 165–180, 2010.

[19] S. Sengan, O. I. Khalaf, S. Priyadarsini, D. K. Sharma, K. Amarendra, and A. A. Hamad, "Smart healthcare security device on medical IoT using raspberry pi," *International Journal of Reliable and Quality E-Healthcare*, vol. 11, no. 3, pp. 1–11, 2022.

[20] S. Quattrini, B. Pampaloni, and M. L. Brandi, "Natural mineral waters: chemical characteristics and health effects," *Clinical cases in mineral and bone metabolism: The Official*

Journal of the Italian Society of Osteoporosis, Mineral Metabolism, and Skeletal Diseases, vol. 13, no. 3, pp. 173–180, 2016.

[21] R. Schäffer, K. Bär, S. Fischer, J. G. Fritsche, and I. Sass, "Mineral, thermal and deep groundwater of Hesse, Germany," *Earth System Science Data*, vol. 13, no. 10, pp. 4847–4860, 2021.

[22] V. Sissakian, N. Al-Ansari, I. Issa, N. Adamo, and S. Knutsson, "Mystery of Mosul dam the most dangerous dam in the world: general geology," *Journal of Earth Sciences and Geotechnical Engineering*, vol. 5, pp. 1–13, 2015.

Review Article

Accuracy and Utility of Preoperative Ultrasound-Guided Axillary Lymph Node Biopsy for Invasive Breast Cancer: A Systematic Review and Meta-Analysis

Yihong Huang ,[1] **Shuo Zheng,**[2] **and Yu Lin**[1]

[1]*Department of Ultrasound, Fuzhou Second Hospital, Fuzhou 350007, China*
[2]*Department of Hepatobiliary Surgery, Fuzhou Second Hospital, Fuzhou 350007, China*

Correspondence should be addressed to Yihong Huang; h18259083272@126.com

Background. With the acceleration of the pace of life and work, the incidence rate of invasive breast cancer is getting higher and higher, and early diagnosis is very important. This study screened and analyzed the published literature on ultrasound-guided biopsy of invasive breast cancer and obtained the accuracy and practicality of preoperative biopsy. *Method.* The four databases were screened for the literature. There was no requirement for the start date of retrieval, and the deadline was July 2, 2022. Two researchers screened the literature, respectively, and included the literature on preoperative ultrasound-guided biopsy and intraoperative and postoperative pathological diagnosis of invasive breast cancer. The diagnostic data included in the literature were extracted and meta-analyzed with RevMan 5.4 software, and the bias risk map, forest map, and summary receiver operating characteristic curves (SROC) were drawn. *Results.* The included 19 studies involved about 18668 patients with invasive breast cancer. The degree of bias of the included literature is low. The distribution range of true positive, false positive, true negative, and false negative in the forest map is large, which may be related to the large difference in the number of patients in each study. Most studies in the SROC curve are at the upper left, indicating that the accuracy of ultrasound-guided axillary biopsy is very high. *Conclusion.* For invasive breast cancer, preoperative ultrasound-guided biopsy can accurately predict staging and grading of breast cancer, which has important reference value for surgery and follow-up treatment.

1. Introduction

Invasive breast cancer is a kind of malignant tumor [1], and no definite therapeutic factors have been found [2]. With the acceleration of the pace of life and work and the increase of pressure, the incidence rate of invasive breast cancer in women has increased year by year [3]. The main symptoms are painless breast tumors [4]. Early invasive breast cancer can be cured by timely treatment after discovery [5]. After the discovery of advanced invasive breast cancer, timely surgical resection, adjuvant radiotherapy and chemotherapy, and targeted drug therapy can effectively control the progress of the disease [6], with a survival rate of 70% or more. The prognosis is related to the stage of cancer [7]. Breast cancer is divided into invasive breast cancer and noninvasive breast cancer [8]. Invasive breast

cancer can invade surrounding tissues and has the ability of distant metastasis [9]. Breast cancer has a high degree of malignancy [10]. According to the World Health Organization (WHO) histological classification of breast tumors [11], invasive breast cancer is divided into special breast cancer and nonspecial breast cancer [12]. Special breast cancer is divided into simple tubular carcinoma, invasive cribriform carcinoma, medullary carcinoma, and mucinous carcinoma [13, 14]. Nonspecific breast cancer is divided into invasive ductal carcinoma and invasive lobular carcinoma [15], of which invasive ductal carcinoma is the most common breast cancer [16, 17], accounting for about 70% to 80% of breast cancer.

At present, no clear therapeutic factors for breast cancer have been found. It is generally believed that pathogenic factors are family history and genetics [18], endogenous

estrogen, exogenous estrogen, radiation exposure, benign breast disease and breast cancer in situ, alcohol intake, caffeine intake through coffee and tea [3], weight gain or overweight, sedentary, diet dominated by meat and sweets, and insufficient vitamin intake [19, 20]. Early detection and treatment are very important [21]. The grading and staging of breast cancer has a significant impact on prognosis [22, 23]. The WHO takes the Nottingham grading system as the standard histological grading system for invasive breast cancer [24, 25]. The evaluation indicators are the proportion of glandular duct formation, nuclear pleomorphism, and mitotic image count [26]. The widely used biopsy method for the diagnosis of breast cancer is ultrasound-guided axillary lymph node biopsy, observing the removed tissue under the microscope to make an accurate diagnosis [27]. In case of breast cancer, surgical resection treatment should be carried out as soon as possible [18]. This study searched all the literature of PubMed, Embase, Cochrane, and Web of Science, screened out the literature related to preoperative ultrasound-guided axillary lymph node biopsy in patients with invasive breast cancer, and at the same time, identified the results of intraoperative and postoperative pathological diagnosis, meta-analyzed the accuracy and practicality of this biopsy, and drew a meaningful conclusion for clinical treatment.

2. Methods

2.1. Literature Search. We searched the database PubMed, Embase, Cochrane Library, and Web of Science with Mesh terms and keywords [28]. The Mesh terms were "invasive breast cancer" and "biopsy," and "invasive breast cancer" includes all the types of invasive breast cancer mentioned above. In order to search the literature as much as possible, we did not take ultrasound-guided biopsy as the subject word. First, we searched all the biopsy literature and then screened and removed the literature that did not use ultrasound-guided biopsy. In PubMed, the Cochrane Library, and the Web of Science database, the search term is ((Invasive breast cancer) OR (Nonspecific invasive breast cancer) OR (Invasive ductal carcinoma) OR (Invasive lobular carcinoma) OR (Special types of invasive breast cancer) OR (Simple tubular carcinoma) OR (Invasive cribriform carcinoma) OR (Medullary carcinoma) OR (Mucinous carcinoma)) AND ((Biopsy) OR (Needle Biopsy) OR (Fine biopsy) OR (Endoscopic Ultrasound-Guided Fine Needle Aspiration) OR (Large-Core Needle Biopsy) OR (core biopsy) OR (Needle Biopsies) OR (Needle Biopsy) OR (Aspiration Biopsy) OR (Aspiration Biopsies) OR (Puncture Biopsy) OR (Puncture Biopsies) OR (Fine-Needle Biopsy)). In the Embase database, the search term is ("invasive breast cancer": ti OR "nonspecific invasive breast cancer": ti OR "invasive ductal carcinoma": ti OR "invasive lobular carcinoma": ti OR "special types of invasive breast cancer": ti OR "simple tubular carcinoma": ti OR "invasive cribriform carcinoma": ti OR "medullary carcinoma": ti OR "mucinous carcinoma": ti) AND (biopsy: ti OR "fine biopsy": ti OR "endoscopic ultrasound-guided fine needle aspiration": ti OR "large-core needle biopsy": ti OR "core biopsy": ti OR

"needle biopsies": ti OR "needle biopsy": ti OR "aspiration biopsy": ti OR "aspiration biopsies": ti OR "puncture biopsy": ti OR "puncture biopsies": ti OR "fine-needle biopsy": ti). The latest retrieval time is July 2, 2022. There is no time limit for the literature in the database, and the retrieval starts from the earliest establishment time of the database. Two independent researchers searched the database, respectively, screened the literature together, and decided to include the analyzed literature through discussion and consultation. The references included in the literature are further screened to determine whether there are missing documents. There are no language restrictions in searching and screening the literature.

2.2. Literature Screening. The retrieved literature was screened according to the inclusion and exclusion criteria. Inclusion criteria were as follows: (1) complete literature can be obtained; (2) the literature mainly describes the diagnosis and treatment of patients with invasive breast cancer; (3) there are preoperative biopsy reports and intraoperative and postoperative pathological diagnosis reports of patients with invasive breast cancer in the literature; (4) preoperative biopsy was axillary lymph node puncture biopsy guided by ultrasound; (5) node puncture biopsy includes ultrasound-guided fine needle aspiration (US-FNA), ultrasound-guided core needle biopsy (US–CNB), and other puncture biopsies; (6) the literature has no language restrictions and no publication time restrictions.

Exclusion criteria were as follows: (1) duplicate literature; (2) review literature; (3) meta-analysis literature; (4) literature on metastasis of primary cancers such as pancreatic cancer, thyroid cancer, liver cancer, gastric cancer, or cervical cancer on the breast and armpit; (5) literature unrelated to preoperative biopsy, such as cryosurgery, chemoradiotherapy, and biopsy technology; (6) unclear description or results and incomplete patient data records.

2.3. Data Collection. Two researchers screened the literature and independently extracted the data from the selected literature using structured data collection tables. We extracted the basic data of patients with invasive breast cancer included in the literature and the basic situation of the literature. The extracted variables include true positive, false positive, true negative, and false negative; that is, the results of intraoperative and postoperative pathological diagnosis are compared with the results of preoperative ultrasound-guided axillary lymph node biopsy, and the former shall prevail to obtain the true and false diagnosis results of the latter. The progress information of postoperative patients was extracted, the information of postoperative complications was extracted, and relevant information was extracted from patients with long-term follow-up records.

2.4. Statistical Analysis. The Review Manager 5.4 (RevMan 5.4) software of the Cochrane Collaboration Network was used to evaluate the bias risk of the included literature. The

specific evaluation contents included the generation of random sequences, allocation concealment, blinding of subjects, blinding of result evaluators, data integrity, and selective reporting. If the opinions assessed by the two researchers were not the same, a third researcher was required to participate in the assessment. RevMan 5.4 software was used to compare the preoperative biopsy results with the intraoperative and postoperative pathological results of all patients with invasive breast cancer included in the literature, and the forest map was made and analyzed. The specificity was studied using 95% confidence interval and calculated using the Mantel– Haenszel random-effect model. Funnel charts were made for all the included literature to visually and clearly reflect the bias.

3. Results

3.1. Literature Search and Screening Results and Basic Information of the Included Literature.

Four databases were searched, and a total of 647 documents were obtained. Among them, there are 135 PubMed, 155 Embase, 30 Cochrane Library, and 327 web of science. Among them, 137 literature reviews were repeated. After reading the title and abstract, combined with the conditions of inclusion and exclusion of documents, 411 documents were eliminated. After intensive reading of the articles, 80 articles were eliminated, and 19 articles were finally included in meta-analysis. The flowchart of literature screening is shown in Figure 1.

The authors, countries, languages, the number of patients with invasive breast cancer, and types of preoperative ultrasound-guided biopsy of 19 literature reviews were extracted and listed. As shown in Table 1, the total number of patients with invasive breast cancer included in the literature was about 18668. All the above work was completed by two researchers.

3.2. Bias-Risk Assessment of Included Articles.

RevMan 5.4 software was used to analyze the bias of 20 included literature reviews. The literature reviews were analyzed from four aspects: patient selection, index test, reference standard, and flow and timing (see Figure 2 for details).

3.3. Forest Plot.

The forest map was drawn with RevMan 5.4 software, and false positive, false negative, true positive, and true negative of 19 studies were counted (see Figure 3 for details).

3.4. SROC Curve.

Using RevMan 5.4 software to draw the SROC curve, it is found that most studies are distributed in the upper left, some of which are close to 1, and only two studies are distributed in the lower left (see Figure 4).

4. Discussion

Invasive breast cancer metastasis to axillary lymph nodes can help determine the stage of invasive breast cancer. Through imaging examination of other parts of the body, if no metastasis is found, it is determined to be in the early stage of breast cancer. Lymph node metastasis in breast cancer is usually stage 2 or more. After the tumor focus of breast cancer metastasizes to ipsilateral axillary lymph nodes, it can still be pushed locally, indicating that it has entered stage 2 breast cancer. Stage 3 of breast cancer will present after tumor foci have metastasized to the ipsilateral axillary lymph nodes. Metastasis develops to supraclavicular lymph node metastasis, and the patient also has distant organ metastasis, which indicates that it is stage 4 of breast cancer. Ultrasound-guided axillary lymph node biopsy can make a more accurate judgment on the staging of breast cancer, so as to guide the treatment method and the surgical resection scope. At present, there are many new biopsy technologies, such as nuclear magnetic resonance-guided biopsy, and ultrasound-guided biopsy is the most widely used and longest used biopsy technology. There is an urgent need to summarize and analyze the accuracy and applicability of this technology to provide guidance for clinical biopsy of breast cancer.

In this study, 19 literature reviews were selected to compare the preoperative ultrasound-guided axillary biopsy of invasive breast cancer with the intraoperative and postoperative pathological results, and we found out the number of false positive, false negative, true positive, and true negative and drew the forest map and SROC curve. The results showed that most of the studies had high sensitivity and specificity, most of the studies were on the upper left of the SROC curve, and some of these studies are close to 1, indicating that preoperative ultrasound biopsy has high diagnostic accuracy and can effectively predict the metastasis of breast cancer.

In recent years, there have been many new techniques for preoperative biopsy of breast cancer. The application of imaging in the diagnosis of breast cancer is becoming more and more mature, especially the application of ultrasound technology. Chung et al. [47] compared the role of ultrasound, CT, MRI, and PET/CT in predicting axillary lymph node metastasis in breast cancer. 1472 patients with invasive breast cancer with ultrasonic staging of lymph nodes were examined by the above nonultrasonic examination. By comparing with the status of biopsy lymph nodes, it is concluded that the accuracy of ultrasound diagnosis of the supraclavicular region, suspicious supraclavicular lymph nodes, and the IM region is more than 93%, and the overall accuracy of other imaging examinations is lower than that of ultrasound. Zhang et al. [48] compared three ultrasound techniques to detect and predict the risk of axillary lymph node (AlN) metastasis of breast invasive ductal carcinoma. They found that when conventional ultrasound (C-US), ultrasonic elastography (UE), and percutaneous contrast-enhanced ultrasound (P-CUES) were combined, their sensitivity, specificity, positive predictive value, and negative predictive value were 94%, 89%, 86%, and 95%, respectively, which were higher than the detection and prediction results.

It can be seen that ultrasound technology itself can accurately and clearly diagnose breast cancer, while biopsy guided by gold standard ultrasound and histopathological analysis can make an accurate diagnosis of malignant invasive breast cancer. Ji et al. [49] evaluated metastasis of
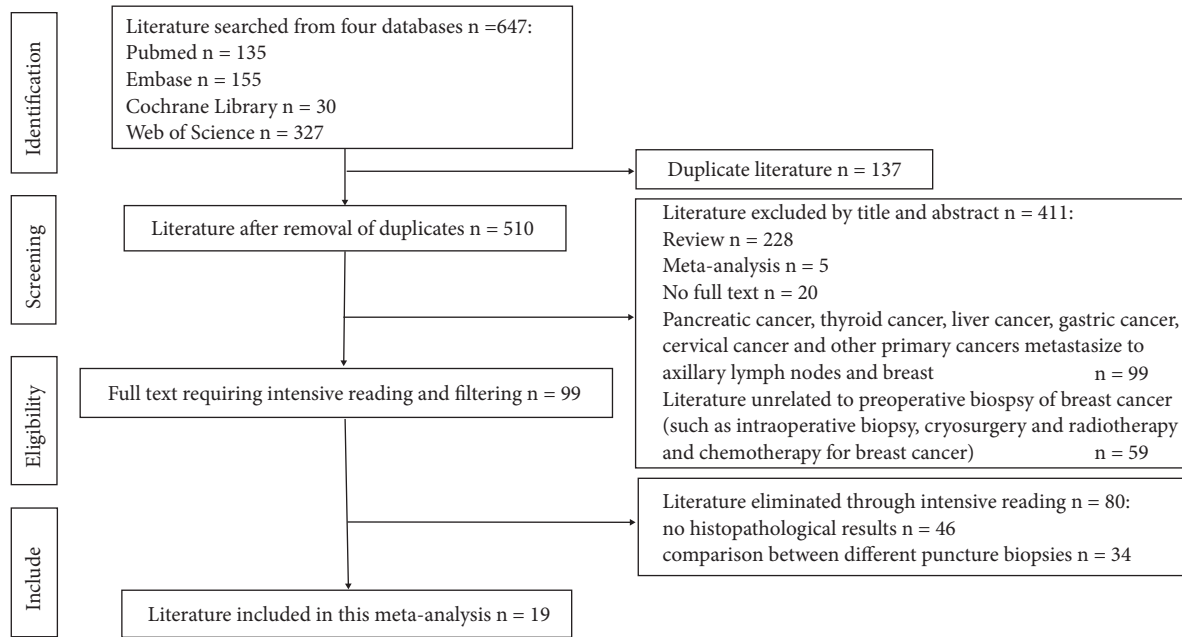
FIGURE 1: The literature screening flowchart. 647 articles were retrieved from four databases, and 22 articles were included in meta-analysis after screening.

TABLE 1: Basic information and the patient information in the literature.

	Country	Language	Number of patients	Preoperative ultrasound-guided biopsy
Osanai [29]	Japan	English	31	US-CNB
Britton et al. [30]	U.K.	English	142	CB, SLN, ALND
Engohan-Aloghe et al. [31]	Belgium	English	71	US, ALND, USG-FNA
Evans and Lyons [32]	U.K.	English	1562	US
Jankowski et al. [33]	France	English	121	SLNB, ALND
Novak et al. [34]	Slovenia	English	102	US-FNAB, AUS
Morrow et al. [35]	U.K.	English	5076	AUS, FNAC
Kim et al. [36]	Korea	English	142	US, US-FNA
Cowher et al. [37]	U.S.A	English	152	AUS
Afzal et al. [38]	Pakistan	English	50	SLNB
Bode and Rissanen [39]	Finland	English	25	US, CNB
Wu et al. [40]	Taiwan, China	English	513	US-CNB
O'Leary [41]	—	English	113	CNB
Rautiainen et al. [42]	Finland	English	54	US-CNB
Park et al. [43]	Korea	English	3124	US-14GCNB
Tahir et al. [44]	U.K.	English	197	US-FNAC
Rautiainen et al. [42]	Finland	English	178	US-FNAB, CNB
Topps et al. [45]	U.K.	English	<417	AUS-FNA, AUS-SNB, AUS-ALND
Zheng et al. [46]	Canada	English	300	US-CB

Note. US-CNB, ultrasound-guided automated percutaneous core needle biopsy. CB, core biopsy. SLN, sentinel lymph node. ALND, axillary lymph node dissection. USG-FNA, ultrasound-guided fine needle aspiration. US, ultrasonography. UNB, core needle biopsy. SLNB, sentinel lymph node biopsy. US-FNAB, ultrasound-guided fine needle aspiration biopsy. AUS, axillary ultrasonography. FNAC, fine needle cytology. US-FNA, ultrasound-guided fine needle aspiration. CNB, core needle biopsy. US-CNB, ultrasound-guided core needle biopsy. US-UNB, ultrasound guided-axillary lymph node core biopsy. US-14GCNB, ultrasound guided-14-gauge core needle biopsy. US-FNAC, ultrasound-guided fine needle aspiration biopsy. US-CNB, ultrasound-guided axillary lymph node core biopsy. AUS-FNA, axillary ultrasound-fine needle aspiration. AUS-SNB, axillary ultrasound-sentinel needle biopsy. AUS-ALND, axillary ultrasound-axillary lymph node dissection. US-CB, ultrasound-guided core biopsy.

breast lymph nodes in breast cancer by ultrasound-guided core real needle biopsy (CNB). The results showed that 131 of the 164 internal mammary lymph nodes treated with CNB were confirmed to be metastasis positive by histopathology, where 8 were negative and 25 were in an unknown state, indicating that ultrasound can accurately detect lymph nodes that may be malignant. Real time ultrasound-guided CNB and fine needle biopsy (FNA) are accurate and valuable techniques to determine the condition of breast cancer. Wahab et al. [50] conducted a meta-analysis on pure flat

(a)

(b)
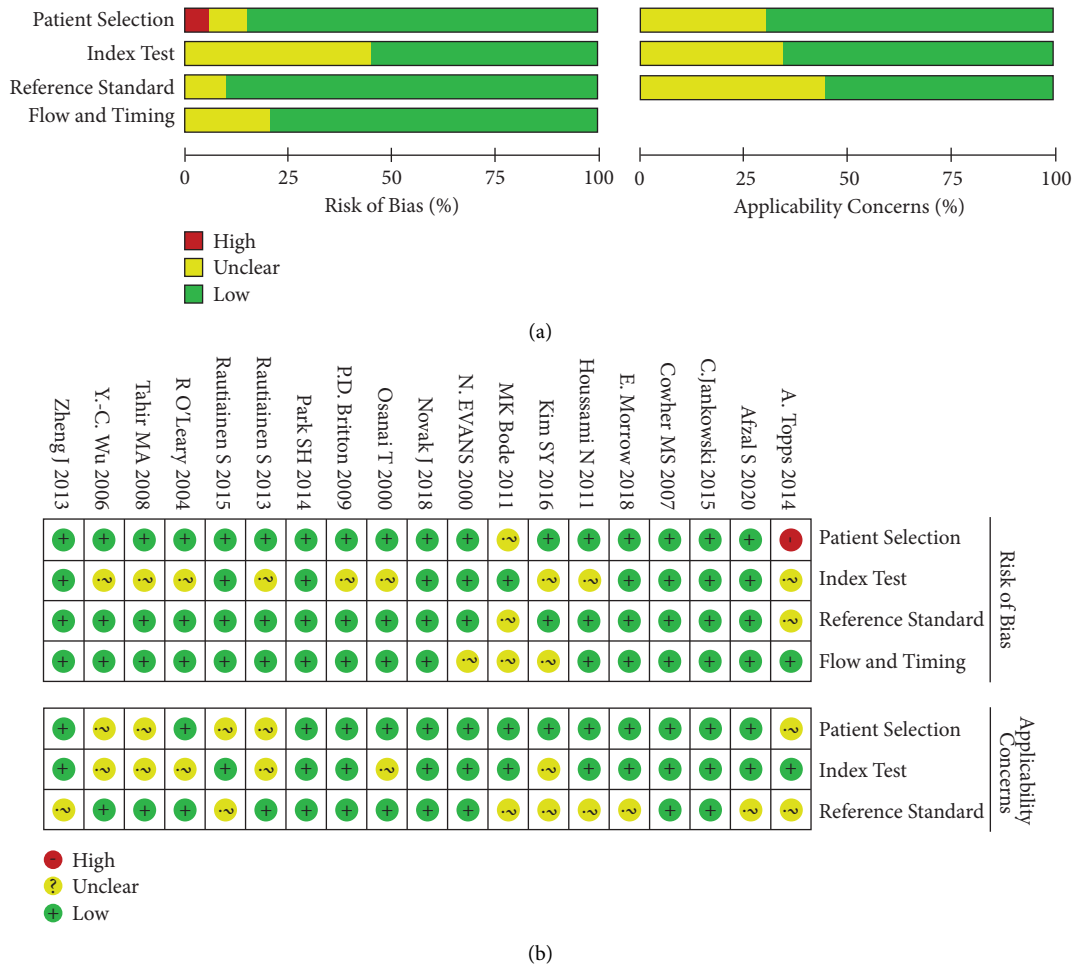
FIGURE 2: The bias of 20 included literature reviews is drawn with RevMan 5.4. (a) Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies. (b) Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
A. Topps 2014	78	0	79	0	0.50 [0.42, 0.58]	Not estimable		
Afzal S 2020	78	0	3	25	0.88 [0.69, 0.97]	1.00 [0.86, 1.00]		
C.Jankowski 2015	28	25	34	34	0.45 [0.32, 0.58]	0.58 [0.44, 0.70]		
Cowher MS 2007	31	3	26	65	0.54 [0.41, 0.68]	0.96 [0.88, 0.99]		
E. Morrow 2018	622	493	3457	123	0.15 [0.14, 0.16]	0.20 [0.17, 0.23]		
Engohan-Aloghe C2009	14	12	12	33	0.54 [0.33, 0.73]	0.73 [0.58, 0.85]		
Kim SY 2016	10	75	8	60	0.56 [0.31, 0.78]	0.44 [0.36, 0.53]		
MK Bode 2011	24	1	0	1	1.00 [0.86, 1.00]	0.50 [0.01, 0.99]		
N. EVANS 2000	14	1	1	0	0.93 [0.68, 1.00]	0.00 [0.00, 0.97]		
Novak J 2018	32	13	14	43	0.70 [0.54, 0.82]	0.77 [0.64, 0.87]		
Osanai T 2000	17	1	1	75	0.94 [0.73, 1.00]	0.99 [0.93, 1.00]		
P.D. Britton 2009	39	41	25	34	0.61 [0.48, 0.73]	0.45 [0.34, 0.57]		
Park SH 2014	60	0	9	0	0.87 [0.77, 0.94]	Not estimable		
R O’Leary 2004	78	0	0	35	1.00 [0.95, 1.00]	1.00 [0.90, 1.00]		
Rautiainen S 2013	45	0	6	99	0.88 [0.76, 0.96]	1.00 [0.96, 1.00]		
Rautiainen S 2015	12	29	0	42	1.00 [0.74, 1.00]	0.59 [0.47, 0.71]		
Tahir MA 2008	8	0	10	20	0.44 [0.22, 0.69]	1.00 [0.83, 1.00]		
Y.-C. Wu 2006	442	0	5	67	0.99 [0.97, 1.00]	1.00 [0.95, 1.00]		
Zheng J 2013	196	0	0	0	1.00 [0.98, 1.00]	Not estimable		

FIGURE 3: Forest plot. Values are shown with 95 percent confidence interval.
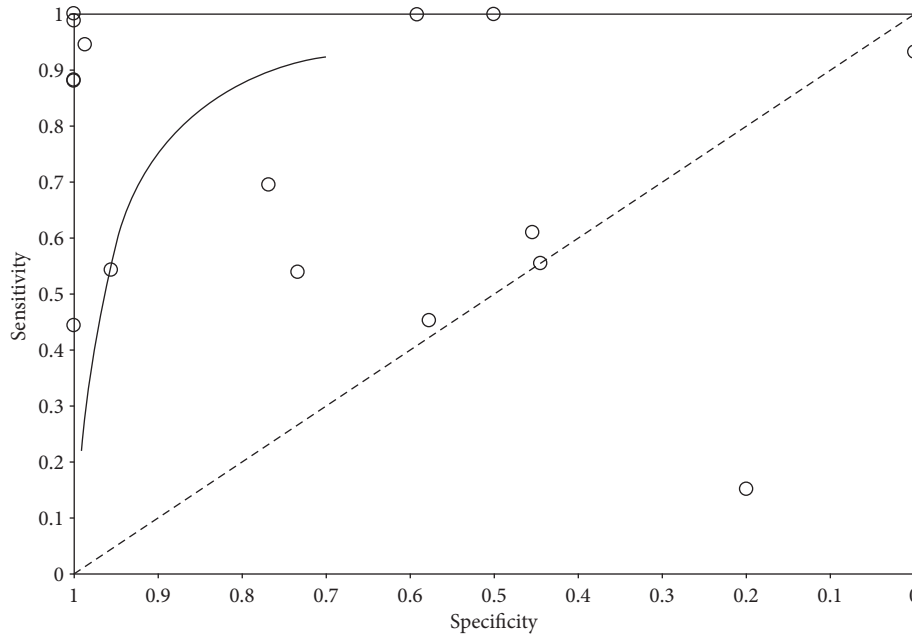
Figure 4: SROC curve.

epithelial atypical (FEA) lesions diagnosed in core needle biopsy (CNB). The results showed that when the combined escalation rate of breast cancer was 5%, pure FEA diagnosed by CNB should be surgically removed. If more than 90% of the targeted calcification was removed by CNB, close imaging follow-up was recommended. Shehata et al. [51] conducted a meta-analysis on the risk of upgrading to malignant tumors after the diagnosis of a lobular tumor by core needle biopsy of the breast. Through reading and summarizing a large number of literature reviews, it was concluded that the risk of upgrading to malignant tumors was less than 45%, and the risk was low. It was speculated that imaging examination was likely to be an alternative method of surgery. Song [52] performed a meta-analysis on the accuracy of targeted axillary lymph node biopsy (TLNB) in breast cancer patients with positive initial lymph nodes. Regression analysis showed that the overlap of the results of targeted and sentinel lymph node biopsy may be related to the identification rate (IFR) and the false negative rate (FNR), while the new technology TLNB has good IFR, low FNR, and high NPV. On the other hand, the relationship between breast cancer and other tumors also deserves attention [53].

In this study, the preoperative ultrasound-guided biopsy of invasive breast cancer was meta-analyzed and compared with the final diagnostic results. There are some limitations in this study that need to be supplemented by subsequent research. The literature searched is still small, so we should search other databases such as clinical trials and national libraries of various countries. It is necessary to use the full-text search function to retrieve the literature related to preoperative ultrasound biopsy. There may be some literature related to relevant content, but it is not the subject content of the literature. This kind of literature should also be carefully read and screened, and there may

be some gains. Ultrasound-guided biopsy is the most widely used. In recent years, MRI-guided biopsy and X-ray interventional biopsy have appeared, which can compare various imaging biopsies and draw clinically meaningful conclusions. Preoperative biopsies of different types of invasive breast cancer can be studied to find the differences, such as the location of invaded axillary lymph nodes, so as to conduct in-depth research on invasive breast cancer [54, 55].

5. Conclusions

For invasive breast cancer, preoperative ultrasound-guided axillary lymph node biopsy can accurately predict the grading and staging of breast cancer, with an accuracy of more than 95%, which can provide a reference for surgery. The histopathological examination of the tumor during and after the operation was highly consistent with the preoperative biopsy, which confirmed that the accuracy of ultrasound-guided biopsy was very high. At present, ultrasound-guided biopsy of breast cancer is the most widely used technology, and other imaging methods cannot compete with it. Preoperative ultrasound-guided biopsy plays a key role in the operation and treatment of breast cancer or even a decisive role.

Data Availability

The data used in this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Acknowledgments

References

[1] K. P. Trayes and S. Cokenakes, "Breast cancer treatment," *American Family Physician*, vol. 104, no. 2, pp. 171–178, 2021.

[2] P. Magesh, S. Thankachan, T. Venkatesh, and P. S. Suresh, "Breast cancer fibroblasts and cross-talk," *Clinica Chimica Acta*, vol. 521, pp. 158–169, 2021.

[3] K. H. Zheng, K. Zhu, J. Wactawski Wende et al., "Caffeine intake from coffee and tea and invasive breast cancer incidence among postmenopausal women in the Women's Health Initiative," *International Journal of Cancer*, vol. 149, no. 12, pp. 2032–2044, 2021.

[4] E. Tarighati, H. Keivan, and H. Mahani, "A review of prognostic and predictive biomarkers in breast cancer," *Clinical and Experimental Medicine*, 2022.

[5] X. Zhang, "A three-mRNA signature associated with pyrimidine metabolism for prognosis of BRCA," *BioMed Research International*, vol. 2022, Article ID 7201963, 2022.

[6] L. Gianni, C. S. Huang, D. Egle et al., "Pathologic complete response (pCR) to neoadjuvant treatment with or without atezolizumab in triple-negative, early high-risk and locally advanced breast cancer: NeoTRIP Michelangelo randomized study," *Annals of Oncology*, vol. 33, no. 5, pp. 534–543, 2022.

[7] P. Chen, X. Zhang, R. Ding et al. "Patient-derived organoids can guide personalized-therapies for patients with advanced breast cancer," *Advanced Science*, vol. 8, no. 22, Article ID e2101176, 2021.

[8] E. Mohammadi, M. Tabatabaei, M. Habibi-Anbouhi, and M. Tafazzoli-Shadpour, "Chemical inhibitor anticancer drugs regulate mechanical properties and cytoskeletal structure of non-invasive and invasive breast cancer cell lines: study of effects of Letrozole, Exemestane, and Everolimus," *Biochemical and Biophysical Research Communications*, vol. 565, pp. 14–20, 2021.

[9] T. Gathani, G. Reeves, J. Broggio, and I. Barnes, "Ethnicity and the tumour characteristics of invasive breast cancer in over 116,500 women in England," *British Journal of Cancer*, vol. 125, no. 4, pp. 611–617, 2021.

[10] G. Zheng, J. Sundquist, K. Sundquist, and J. Ji, "Family history of breast cancer as a second primary malignancy in relatives: a nationwide cohort study," *BMC Cancer*, vol. 21, no. 1, p. 1210, 2021.

[11] A. Lebeau, "[Updated WHO classification of tumors of the breast]," *Pathologe, Der*, vol. 42, no. Suppl 2, pp. 155–159, 2021.

[12] J. Yu, J. Wu, O. Huang et al., "Distribution and influence of the 21-gene recurrence score on chemotherapy decision-making in special type of breast cancer," *American journal of cancer research*, vol. 11, no. 12, pp. 6188–6199, 2021.

[13] R. A. McCart, "Invasive lobular carcinoma of the breast: the increasing importance of this special subtype," *Breast Cancer Research*, vol. 23, no. 1, p. 6, 2021.

[14] E. J. Watkins, "Overview of breast cancer," *Journal of the American Academy of Physician Assistants*, vol. 32, no. 10, pp. 13–17, 2019.

[15] C. M. Zhou, Q. Xue, Y. Wang, J. Tong, M. Ji, and J.-J. Yang, "Machine learning to predict the cancer-specific mortality of patients with primary non-metastatic invasive breast cancer," *Surgery Today*, vol. 51, no. 5, pp. 756–763, 2021.

[16] N. Eshghifar, F. Rouhollah, N. Barikrow, F. Pouresmaeili, and M. Taheri, "The role of long noncoding RNAs in patients with Luminal A invasive breast ductal carcinoma," *Pathology, Research & Practice*, vol. 227, Article ID 153645, 2021.

[17] F. Kanavati and M. Tsuneki, "Breast invasive ductal carcinoma classification on whole slide images with weakly-supervised and transfer learning," *Cancers*, vol. 13, no. 21, 2021.

[18] J. J. Moya, A. Moazzez, J. J. Ozao-Choy, and C. Dauphine, "Patients with invasive breast cancer who refuse treatment: an analysis of associated factors and impact on survival," *The American Surgeon*, vol. 87, no. 10, pp. 1627–1632, 2021.

[19] Y. Min, X. Liu, D. Hu et al., "Risk factors, prognostic factors, and nomogram for distant metastasis in breast cancer patients without lymph node metastasis," *Frontiers in Endocrinology*, vol. 12, Article ID 771226, 2021.

[20] F. Ntirenganya, J. D. Twagirumukiza, G. Bucyibaruta, B. Rugwizangoga, and S. Rulisa, "Premenopausal breast cancer risk factors and associations with molecular subtypes: a case-control study," *International journal of breast cancer*, vol. 2021, Article ID 5560559, 2021.

[21] R. H. Hester, G. N. Lim, and B. Lim, "Inflammatory breast cancer: early recognition and diagnosis is critical," *American Journal of Obstetrics and Gynecology*, vol. 225, no. 4, pp. 392–396, 2021.

[22] T. Friedl, T. Fehm, V. Müller et al. "Prognosis of patients with early breast cancer receiving 5 Years vs. 2 Years of adjuvant bisphosphonate treatment," *JAMA Oncology*, vol. 7, no. 8, pp. 1149–1157, 2021.

[23] B. K. Kim, J. M. Ryu, S. J. Oh et al., "Comparison of clinicopathological characteristics and prognosis in breast cancer patients with different Breast Imaging Reporting and Data System categories," *Annals of Surgical Treatment and Research*, vol. 101, no. 3, pp. 131–139, 2021.

[24] S. Mantrala, "Concordance in breast cancer grading by artificial intelligence on whole slide images compares with a multi-institutional cohort of breast pathologists," *Archives of Pathology & Laboratory Medicine*, 2022.

[25] S. Kurozumi, K. Kaira, H. Matsumoto et al., "Association of L-type amino acid transporter 1 (LAT1) with the immune system and prognosis in invasive breast cancer," *Scientific Reports*, vol. 12, no. 1, p. 2742, 2022.

[26] D. Santucci, E. Faiella, A. Calabrese et al., "On the additional information provided by 3T-MRI ADC in predicting tumor cellularity and microscopic behavior," *Cancers*, vol. 13, no. 20, 2021.

[27] K. Astvatsaturyan, A. Bose, and S. Bose, "Is ultrasound-guided fine needle aspiration biopsy of axillary lymph nodes a viable alternative to sentinel lymph node biopsy?" *Diagnostic Cytopathology*, vol. 49, no. 10, pp. 1099–1109, 2021.

[28] D. Wang, Y.-L. Li, D. Qiu, and S.-Y. Xiao, "Factors influencing paternal postpartum depression: a systematic review and meta-analysis," *Journal of Affective Disorders*, vol. 293, pp. 51–63, 2021.

[29] T. Osanai, "Ultrasound-guided core needle biopsy for breast cancer: preliminary report," *Japanese Journal of Clinical Oncology*, vol. 30, no. 2, pp. 65–67, 2000.

[30] P. D. Britton, A. Goud, S. Godward et al., "Use of ultrasound-guided axillary node core biopsy in staging of early breast cancer," *European Radiology*, vol. 19, no. 3, pp. 561–569, 2009.

[31] C. Engohan-Aloghe, N. Hottat, and J. C. Noel, "Accuracy of lymph nodes cell block preparation according to ultrasound features in preoperative staging of breast cancer," *Diagnostic Cytopathology*, vol. 38, no. 1, pp. 5–8, 2010.

[32] N. Evans and K. Lyons, "The use of ultrasound in the diagnosis of invasive lobular carcinoma of the breast less than 10 mm in size," *Clinical Radiology*, vol. 55, no. 4, pp. 261–263, 2000.

[33] C. Jankowski, D. Hudry, D. Vaillant et al., "Evaluation of axillary involvement by ultrasound-guided lymph node biopsy: a prospective study," *Gynecologie Obstetrique & Fertilite*, vol. 43, no. 6, pp. 431–436, 2015.

[34] J. Novak, N. Besic, R. Dzodic, B. Gazic, and A. Vogrin, "Pre-operative and intra-operative detection of axillary lymph node metastases in 108 patients with invasive lobular breast cancer undergoing mastectomy," *BMC Cancer*, vol. 18, no. 1, p. 137, 2018.

[35] E. Morrow, A. Lannigan, J. Doughty et al., "Population-based study of the sensitivity of axillary ultrasound imaging in the preoperative staging of node-positive invasive lobular carcinoma of the breast," *British Journal of Surgery*, vol. 105, no. 8, pp. 987–995, 2018.

[36] S. Y. Kim, E.-K. Kim, H. J. Moon, J. H. Yoon, and M. J. Kim, "Is pre-operative axillary staging with ultrasound and ultrasound-guided fine-needle aspiration reliable in invasive lobular carcinoma of the breast?" *Ultrasound in Medicine and Biology*, vol. 42, no. 6, pp. 1263–1272, 2016.

[37] M. S. Cowher, K. M. Erb, W. Poller, and T. B. Julian, "Correlation of the use of axillary ultrasound and lymph node needle biopsy with surgical lymph node pathology in patients with invasive breast cancer," *The American Journal of Surgery*, vol. 196, no. 5, pp. 756–759, 2008.

[38] S. Afzal, I. Masroor, A. Munir, R. Idress, P. Khan, and S. Khan, "Preoperative ultrasound-guided core biopsy of axillary nodes for staging of clinically negative axilla in breast cancer patients - a pilot study," *Cureus*, vol. 12, no. 1, p. e6718, 2020.

[39] M. K. Bode and T. Rissanen, "Imaging findings and accuracy of core needle biopsy in mucinous carcinoma of the breast," *Acta Radiologica*, vol. 52, no. 2, pp. 128–133, 2011.

[40] Y. C. Wu, D. R. Kuo, and S. J. Kuo, "Personal experience of ultrasound-guided 14-gauge core biopsy of breast tumor," *European Journal of Surgical Oncology*, vol. 32, no. 7, pp. 715–718, 2006.

[41] R. O'Leary, "Agreement between preoperative core needle biopsy and postoperative invasive breast cancer histopathology is not dependent on the amount of clinical material obtained," *Journal of Clinical Pathology*, vol. 57, no. 2, pp. 193–195, 2004.

[42] S. Rautiainen, A. Masarwah, M. Sudah et al., "Axillary lymph node biopsy in newly diagnosed invasive breast cancer: comparative accuracy of fine-needle aspiration biopsy versus core-needle biopsy," *Radiology*, vol. 269, no. 1, pp. 54–60, 2013.

[43] S. H. Park, M. J. Kim, S. J. Kim, and E.-K. Kim, "Ductal carcinoma in situ diagnosed using an ultrasound-guided 14-gauge core needle biopsy of breast masses: can underestimation be predicted preoperatively?" *Ultrasonography*, vol. 33, no. 2, pp. 128–135, 2014.

[44] M. Tahir, K. A. Osman, J. Shabbir et al., "Preoperative axillary staging in breast CancerSaving time and resources," *Breast Journal*, vol. 14, no. 4, pp. 369–371, 2008.

[45] A. Topps, V. Clay, M. Absar et al., "The sensitivity of pre-operative axillary staging in breast cancer: comparison of invasive lobular and ductal carcinoma," *European Journal of Surgical Oncology*, vol. 40, no. 7, pp. 813–817, 2014.

[46] J. Zheng, T. Alsaadi, J. Blaichman et al., "Invasive ductal carcinoma of the breast: correlation between tumor grade determined by ultrasound-guided core biopsy and surgical pathology," *American Journal of Roentgenology*, vol. 200, no. 1, pp. W71–W74, 2013.

[47] H. L. Chung, K. Shin, J. Sun, and J. W. T. Leung, "Extra-axillary nodal metastases in breast cancer: comparison of ultrasound, MRI, PET/CT, and CT," *Clinical Imaging*, vol. 79, pp. 113–118, 2021.

[48] Q. Zhang, E. A. Agyekum, L. Zhu et al., "Clinical value of three combined Ultrasonography modalities in predicting the risk of metastasis to axillary lymph nodes in breast invasive ductal carcinoma," *Frontiers in Oncology*, vol. 11, Article ID 715097, 2021.

[49] X. Ji, D. Li, D. Gao et al., "Value of ultrasound-guided biopsy in evaluating internal mammary lymph node metastases in breast cancer," *Clinical Breast Cancer*, vol. 21, no. 6, pp. 532–538, 2021.

[50] R. A. Wahab, S.-J. Lee, M. E. Mulligan, B. Zhang, and M. C. Mahoney, "Upgrade rate of pure flat epithelial atypia diagnosed at core needle biopsy: a systematic review and meta-analysis," *Radiology: Imaging Cancer*, vol. 3, no. 1, Article ID e200116, 2021.

[51] M. N. Shehata, H. Rahbar, M. R. Flanagan et al., "Risk for upgrade to malignancy after breast core needle biopsy diagnosis of lobular neoplasia: a systematic review and meta-analysis," *Journal of the American College of Radiology*, vol. 17, no. 10, pp. 1207–1219, 2020.

[52] Y. X. Song, "Diagnostic accuracy of de-escalated surgical procedure in axilla for node-positive breast cancer patients treated with neoadjuvant systemic therapy: a systematic review and meta-analysis," *Cancer Medicine*, 2022.

[53] X. Cheng, Q. Chen, and P. Sun, "Natural phytochemicals that affect autophagy in the treatment of oral diseases and infections: a review," *Frontiers in Pharmacology*, vol. 13, Article ID 970596, 2022.

[54] N. Houssami, S. Ciatto, R. M. Turner, H. S. Cody, and P. Macaskill, "Preoperative ultrasound-guided needle biopsy of axillary nodes in invasive breast cancer," *Annals of Surgery*, vol. 254, no. 2, pp. 243–251, 2011.

[55] S. Rautiainen, M. Sudah, S. Joukainen, R. Sironen, R. Vanninen, and A. Sutela, "Contrast-enhanced ultrasound -guided axillary lymph node core biopsy: diagnostic accuracy in preoperative staging of invasive breast cancer," *European Journal of Radiology*, vol. 84, no. 11, pp. 2130–2136, 2015.

Research Article

Computational Analysis of Influencing Factors and Multiple Scoring Systems of Stone Clearance Rate after Flexible Ureteroscopic Lithotripsy

Lei Xia ⓘ, Hanqing Xuan ⓘ, Yang Cao ⓘ, Zhebin Du ⓘ, Hai Zhong ⓘ, and Qi Chen ⓘ

Department of Urology, Ren Ji Hospital, School of Medicine Shanghai Jiaotong University, Shanghai 200127, China

Correspondence should be addressed to Qi Chen; chenqi@mjc-edu.cn

Our research aims at the analysis of various stone scoring systems which are referred to as STONE scoring system (SSS) in this study. GUY's scoring system and RUSS scoring system (RSS) are utilized to predict stone-free status (SFS) after surgery and problems after percutaneous nephrolithotomy (PCNL) for harder stones. The data of 68 patients with renal calculi who received FURL in Ren Ji Hospital from Jan 2020 to Mar 2021 are collected as the study subjects. There were 44 male and 24 female patients, with an average age of 55.6 ± 11.4 years. Reliability analysis of related influencing factors (IF) of stone clearance rate (SCR) and multiple scoring systems after flexible ureteroscopic lithotripsy (FURL) was performed. Relevant factors with statistical significance for postoperative SCR were selected for logistic regression analysis (RA). According to the SSS score, GSS classification, and RUSS score, the SCR after FURL was statistically analyzed. The results showed that the *P* values corresponding to stone position (lower caliceal), cumulative stone diameter (CSD), urinary tract infection, and external physical vibration lithecbole (EPVL) were less than 0.05. The area under the ROC curve of RUSS score, SSS score, and GSS grading was 0.932, 0.841, and 0.533, respectively. The main IF of SCR after FURL were stone location (lower caliceal), CSD, urinary tract infection, and EPVL. The RUSS score system was the best in the evaluation of SCR after FURL. In the previous research, the score systems such as CROES (CRS), SSS, S-ReS, C, and GSS for the prediction of SFS were compared. In our analysis, we have compared the RUSS scoring system which has proven to be giving better results as compared to SSS and GSS. We also performed the regression analysis and found that the stone location shows the strongest correlation of all the other factors for stone clearing rate.

1. Introduction

Urinary calculi are urinary system disease that is commonly found. The clinical study data show that the incidence percentage is between 1% and 5% [1]. According to epidemiological surveys, the disease has shown a significant upward trend in China and even worldwide in the past decade [2]. Generally speaking, urinary calculi have become a nonnegligible disease threatening human health. In recent years, the detection rate of urinary calculi has been increasing, and a variety of treatment methods have gradually emerged. However, according to the location and size of stones, the general renal calculi require surgical treatment [3]. At present, the common surgical methods in clinical practice include extracorporeal shock wave lithotripsy

(ESWL), percutaneous nephrolithotomy (PCNL), and flexible ureteroscopic lithotripsy (FURL) [4]. PCNL is the medical procedure to remove the calculi. In this medical procedure, a tube is inserted into the kidney through the bladder. The kidney stones are broken. Drainage tubes are used to remove the stones. PSNL is used if the stone is too hard to break. The SFR for PCNL varies in between 85 and 93%. Some of the disadvantages include bleeding, injury to pleura, injury to surrounding organs, and infection.

ESWL is generally used when the kidney stone measures from 2 to 2.5 cms. The procedure uses high-intensity acoustic waves. It is an out-patient procedure. ESWL treatment has a stone removal rate that ranges from 50% to 75%. PSNL shows 85 and 93% SCR. Lower the SCR, lower the success rate of the medical procedure. Although ESWL

has high safety, the stone clearance effect is not good [5]. FURL has numerous advantages such as being minimally invasive, flexible, and repeatable, so it has been gradually widely used in the treatment of upper urinary calculi in recent years [6].

Flexible ureteroscope can retrogradely enter the urinary tract through the natural orifice and can visually and clearly observe the lesions in the urinary tract through the visual system and lithotripsy system [7]. This method has high reproducibility during treatment, so the method also has a better therapeutic effect on special populations such as patients with renal calyceal calculi, people with bleeding tendency, people with excessive obesity, people with renal anatomical malformations, and pregnant women [8, 9]. Although FURL has many advantages, it is affected by many factors, so FURL has a large difference for SCR. It is found that the factors affecting the SCR can be summarized as follows: (1) stone: such as stone diameter and location; (2) patient: such as patient weight, whether there is hydronephrosis before treatment; (3) surgeon: the difference in the proficiency of the surgeon will lead to the difference in SCR, and the same surgeon performing stone clearance surgery at different times will also lead to the difference in SCR [10–12]. At present, most of the relevant domestic studies focus on the safety and efficacy of surgery, and there are few studies on the IF of SCR.

Clinically, surgical treatment needs a simple, effective, and accurate method to assess the postoperative SCR in order to select the best treatment. Therefore, how to predict postoperative SCR accurately and effectively has always been the focus of clinical research. At present, several clinically recognized scoring systems are as follows: (1) GSS stone grading is a relatively simple scoring system, which can accurately and objectively reflect the SCR, but it does not explain some details; for example, it does not distinguish multiple stones in different calyces and the same calyx and does not include spinal malformations. However, these patients are at high risk of renal calculi and should be included; (2) RUSS stone scoring system. Relevant clinical studies have shown that this method is suitable for all patients with FURL renal calculi; (3) SSS renal calculi scoring system. Existing studies show that this scoring system can efficiently predict the SCR of nephroscopic lithotripsy [13]. Each scoring system has its advantages and disadvantages. At present, the relevant studies mostly focus on the clinical application of each scoring system, and there are few reports on the comparative study of multiple scoring systems. In [14], the authors study predictive scores for SFR after flexible ureterorenoscopy (FURS). All four scores (SSS score, RUSS, ReSC, and Ito's nomogram) could predict SFR after FURS. In [15], the researchers assess the effect of pelvicalyceal anatomy on stone clearance in the cases with remnant fragments after flexible ureteroscopy. In [16], the authors compare the scoring systems—nephrolithometry, GSS, SSS, and CRS nomogram. They have developed a risk group stratification after assessing the accuracy in prediction for SFR and other variables concerned after surgical operation. In [17], the authors compare the different scoring systems. The pros and cons are elaborated on using different scoring systems. In [18], the authors analyze the anatomical and radiological on stone clearance. They find no statistical significance in BMI, size of the stone, and lower calyx on SCR post-ESWL.

Patients suffering from calculi who underwent FURL in Ren Ji Hospital were selected as the study subjects. The relevant clinical information of the patients was retrospectively studied, and the IF of SCR were analyzed using logistics univariate and multivariate screening. The SCR after FURL was also statistically analyzed according to the SSS score, GSS grading, and RUSS score, to analyze the reliability of different scoring methods. This will provide the basis for the diagnosis and treatment of related diseases.

Major highlights of the paper are as follows:

(i) Analysis of various methods for predictive scores for SFR

(ii) Comparing the predictive scores of the RUSS scoring system, SSS, and GSS

(iii) Identification of the correlation between different factors influencing SCR

2. Methodology and Materials

2.1. Study Samples. A total of 68 patients with renal calculi who received FURL from Jan-2020 to Mar-2021, in Ren Ji Hospital, were selected for the study purpose. There were 44 male patients and 24 female, with a mean age of 55.6 ± 11.4 years. Inclusion criteria were as follows: the patients' clinical data were complete, and the whole process was treated in Ren Ji Hospital, and primary lithotripsy was successful. Single surgeon has handled all the treatments. Surgeon is having more than 2 years of experience in flexible ureteroscopic surgery. Exclusion criteria were as follows: patients with preoperative severe renal insufficiency; patients with severe hydronephrosis; patients who underwent two or more flexible ureteroscopic surgeries; patients with sponge kidney and renal calculi; patients with ureteral stricture. The patients were split into stone clearance group (43 cases) and residual stone group (25 cases).

SCR is the ratio of the number of patients from whom the stone was removed completely to the total patients who received the treatment.

Patient signed an informed consent form, and the experiment met medical ethics requirements.

2.2. Surgical Methods. The patients suffering from urinary tract infection (UTI) were given sensitive antibiotics before the operation, and the operation was performed after the control of urinary tract infection. All patients were given prophylactic antibiotics 0.5–2 hours before the operation to avoid or reduce serious complications such as urosepsis.

After general anesthesia, the lithotomy position was taken, and Wolf F8/9.8 rigid ureteroscope was put in through the urethra, upward along the ureteral orifice of the affected side until the pelvis or upper ureter, and 0.035 mm nickel-titanium wire was inserted. The F12/14 flexible ureteroscope (High Quality Urology Single-Use Flexible

Video Ureteroscope Digital Disposable Urethroscope, Guangzhou Lety Medical, Ltd.) sheath was introduced with the guide wire inserted into the below UPJ or ureteral calculi, the Olympus flexible ureteroscope (URF-V, F8.5-9.9) was inserted along the sheath, and the holmium laser fiber was introduced. The fiber diameter was $200\,\mu$m, and the crushing of stones energy was set to $(0.6$–$1.2)\,J \times (16$–$20)\,Hz$, and the maximum diameter of stones needed to be crushed <2 mm. At the end of the operation, each calyx of the kidney was examined again to ensure that no large stones remained. F5 ureteral stent was put in with the guide wire, and the catheter was indwelled. KUB was reexamined on the first day after FURL to evaluate the crushing of stones and the position of the ureteral stent and to guide the patient's body position according to the distribution of stones. Postoperative expulsion was carried out by referring to the method of Xu Changbao et al. on external physical vibration lithecbole (EPVL).

2.3. Measurement Method.
There are different CT-based measurement methods to calculate stone volume. CT-based 3D-reconstructed algorithm, threshold-based methods, noncontrast helical computed tomography, and manual stone size measurements are some of the methods. [14–16].

Slice computed tomography scanner medical CT scan machine was used to scan the images.

The complete medical history data as well as imaging data such as laboratory tests and CT of each patient were used as evaluation objects. Indicators of patient and stone characteristics as independent variables were collected, and the SCR in one-stage surgery was also calculated. The relevant variables were measured according to RUSS, SSS, and GSS. The average CT value of the stone was measured according to the three areas of the maximum cross section of the stone at three levels by CT scanning, and the CT values (HU) of the core, edge, and the position between the two were recorded, respectively. The average CT value of each stone was calculated (staghorn stones could be measured by this method). The average of multiple stones was calculated. IPA is the minimum angle between the long axis of the lower calyx and the long axis of the ureteropelvic.

Stone size grading: the stones were grouped according to their maximum diameter in the CT cross section, which were divided into diameter less than 1 cm group, diameter 1-2 cm group, and diameter more than 2 cm group. All measurement data were recorded after consensus by the 2 observers.

2.4. Scoring Systems.
GSS stone grading criteria: it was put forward by Thomas et al. in 2011 depending on the structure and stones distribution of renal pelvis and calyces. The GSS was developed after analyzing the published data review and expert opinion. Iterative testing was done. It predicts the post-PCNL SFR with great accuracy. It shows reproducibility and is easy to use. GSS score is found to be significant in predicting the SFR ($P = 0.01$) independently [17]. The method is divided into four levels, and the detailed grading standards and contents are shown in Figure 1.

RUSS was first presented by Resorlu et al. in 2012. The SCR after FURL can be predicted according to the score. The high score signifies low SCR. The detailed scoring criteria are given in Figure 2.

SSS for renal calculi: the scoring system includes five items, and the postoperative SCR of PCNL can be predicted based on the score. The detailed scoring criteria are shown in Table 1.

2.5. Statistical Method.
SPSS 22.0 is utilized in this medical research study for analysis. The data were expressed as mean ± standard deviation ($\overline{x} \pm s$), and for comparison among the groups, t-test was used. Variance analysis was adopted for comparison within the group; χ^2 was used for counting data; $P < 0.05$ shows the difference that is statistically significant. The single factor analysis was carried out using univariate regression analysis. In univariate linear regression, we identify the correlation between single independent variable and one dependent variable. The multifactor analysis was carried out using multivariate RA. In this method, we identify the correlation between single dependent variable and multiple independent variables.

3. Results

3.1. Univariate Regression Analysis Results.
We examined the CT scan reports of patients with renal calculi who received FURL from Jan-2020 to Mar-2021, in Ren Ji Hospital, were selected for the study purpose. Male patients were 44 in number, and female patients were 24 in number. Average age is 55.6 ± 11.4 years.

The univariate analysis is carried out for the IF of the stone clearing rate for the stone clearing group. The IF of the stone clearing rate were analyzed by univariate analysis, and the results were shown in Figure 3. P values corresponding to the location of stones, the cumulative stone diameter (CSD), urinary tract infection, EPVL, hydronephrosis, and other factors are all less than 0.05 indicating that these variables are statistically significant in the univariate RA of the IF of SCR.

The following Table 2 shows the results of univariate RA on IF of SCR.

The following Figure 3 shows the results of univariate RA on IF of SCR.

3.2. Results of Multivariate Regression Analysis.
The IF of the stone clearing rate were analyzed by multiple factors. Figure 4 indicates that the P values corresponding to stone position (lower calyx), stone diameter, urinary tract infection, and EPVL are less than 0.05, with statistical significance, which were finally included in the model. It can be concluded that the location of stones (lower calyx), the maximum diameter of stones (20–30 mm), urinary tract infection, and EPVL are the main factors affecting the SCR after FURL, and the correlation with the location of stones is the strongest.

The following Table 3 shows the multivariate RA results of IF of SCR.
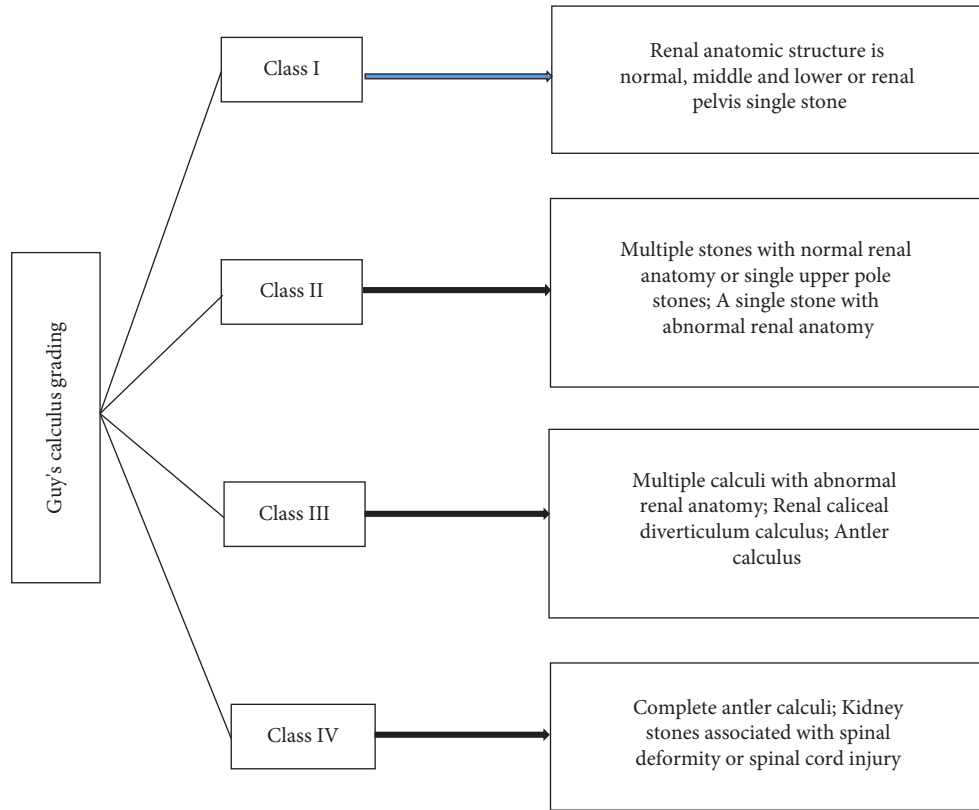
Figure 1: GSS stone grading.
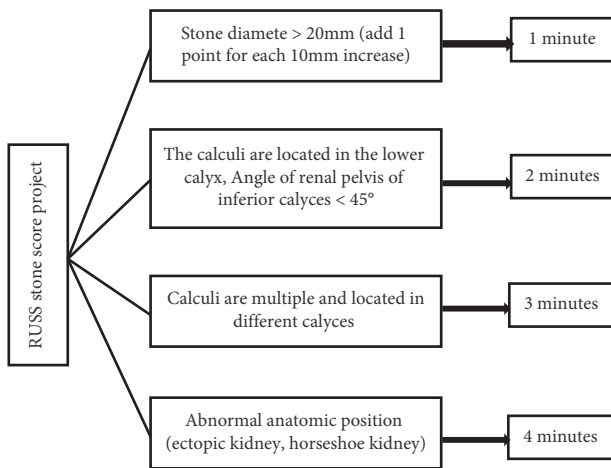
Figure 2: RUSS stone scoring system.

The following Figure 4 shows the multivariate RA results of IF of SCR.

3.3. Comparison of SCR of RUSS Stone Scoring System. SCR comparison results of the RUSS stone scoring system are shown in Figure 5. The proportions of patients with RUSS scores of 0, 1, 2, 3, 4, and 5 in the stone clearance group were 55%, 36%, 36%, 1.8%, 1.8%, and 1.8%, respectively; the proportion of patients with RUSS scores of 0, 1, 2, 3, 4, and 5 in the residual stone group was 12%, 24%, 27%, 21%, 9%, and

6%, respectively. The score of RUSS in the stone clearance group was 0.57 ± 0.44, and that in the residual stone group was 2.11 ± 1.62.

Comparison of SCR of the RUSS stone scoring system is shown in Figure 5.

3.4. Comparison of SCR of STONE. SCR comparison results of SSS were shown in Table 2. SSS stone score in the stone clearance group was 4.33 ± 1.13, the SSS stone score in the stone residual group was 8.26 ± 2.01, and the value of P indicates a statistically significant difference between the two groups. Table 4 shows the comparison of SCR of SSS.

3.5. GSS Stone-Clearance Rate Comparison Results. SCR comparison results of GSS were given in Figure 6. The proportion of patients with GSS grade One, Two, Three, and Four in the stone clearance group was 11%, 45%, 62%, and 0%, respectively; the proportion of patients with GSS grade One, Two, Three, and Four in the stone residual group was 3%, 30%, 64%, and 3%, respectively. Comparison results of GSS are shown in Figure 6.

3.6. ROC Curve Analysis of Each Scoring System for FURL SCR. The ROC curve analysis results of each scoring system for FURL SCR are shown in Figure 7. The ROC curve of each scoring system was 0.932, 0.841, and 0.533, respectively, and the order was RUSS > SSS > GSS. The performance of RUSS

TABLE 1: SSS for renal calculi.

S: stone size, the length × width of the largest cross-sectional area of stone in CT plain scan = cross-sectional area (mm²)	1 points 0~399 2 points: 400~799 3 points: 800~1599 4 points: ≥1600
T: tract length, the distance measured from the center of stone to the skin on the cross section of CT (horizontal line, 45° line, and vertical line).	1 point: ≤100 mm 2 points: >100 mm
O: obstruction, degree of hydronephrosis	1 point: no obstruction or mild hydronephrosis 2 points: moderate or severe hydronephrosis
N: number of involved calices	1 point: 1 renal caliceal involvement 2 points: 2–3 renal calices involvement 3 points: complete staghorn calculi
E: essence or stone density	1 point: ≤950HU 2 points: >950HU
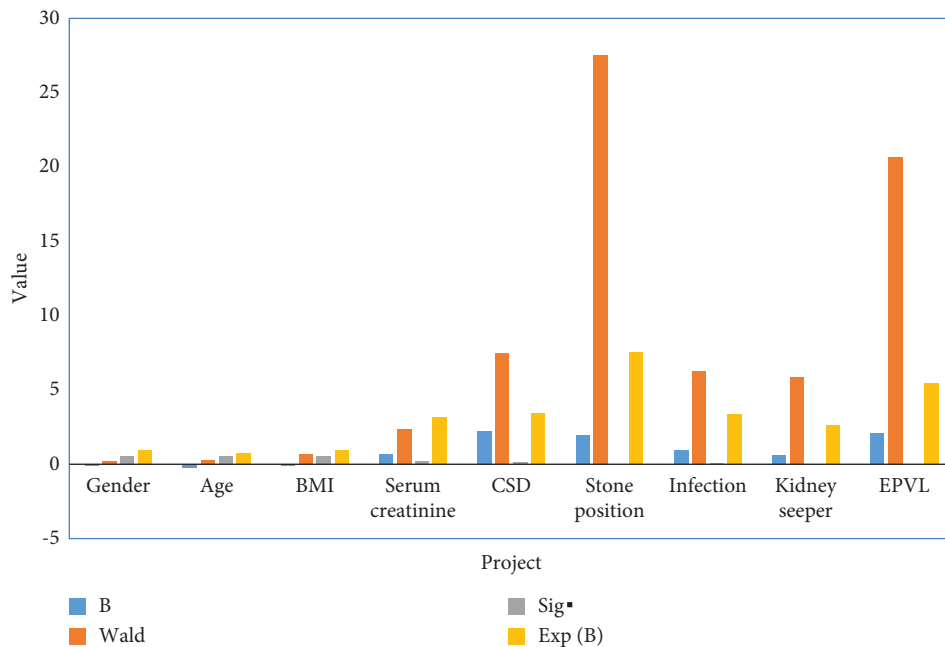
FIGURE 3: Results of univariate RA on IF of SCR. Note: CSD: cumulative stone diameter; EPVL: external physical vibration lithecbole.

TABLE 2: Results of univariate RA on IF of SCR.

	B	Wald	Sig■	Exp (*B*)
Gender	−0.101	0.215	0.541	0.981
Age	−0.241	0.293	0.536	0.785
BMI	−0.143	0.684	0.539	0.932
Serum creatinine	0.657	2.361	0.233	3.149
CSD	2.237	7.459	0.113	3.459
Stone position	1.934	27.53	0	7.532
Infection	0.925	6.303	0.048	3.392
Kidney seeper	0.603	5.882	0.0361	2.639
EPVL	2.077	20.657	0	5.447

scoring system was the best, followed by SSS and GSS. ROC curve analysis of FURL SCR by each scoring system is shown in Figure 7.

4. Discussion

Urinary calculi are one of the general ailments of the urinary system, and the incidence of this disease in China ranges from 5% to 10%, which shows an increasing trend year by year [18]. This disease is also found amongst the children now a days.

New treatments and corresponding grading systems have also been developed. At present, the main treatment methods for the urinary system include ESWL, PCNL, and FURL. URL is currently widely used in clinical practice. FURL is widely adopted due to its characteristics of less trauma, high efficiency, and wide indications has become the treatment of option for renal calculi of 1-2 cm in diameter and is considered to be the best alternative for patients who cannot undergo PCNL. We can gauge the effectiveness of
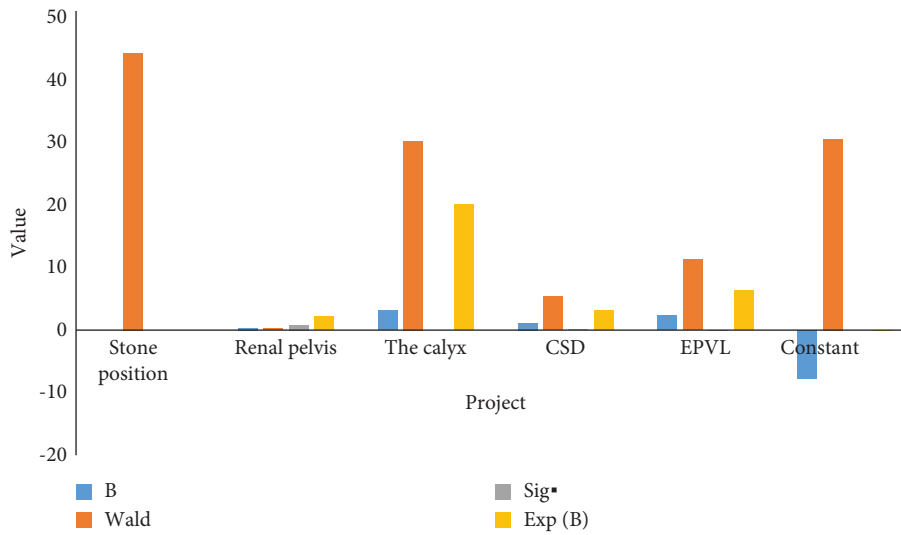
Figure 4: Multivariate RA results of IF of SCR.

Table 3: Multivariate RA results of IF of SCR.

	B	Wald	Sig■	Exp (B)
Stone position	0	44.36	0	0
Renal pelvis	0.361	0.286	0.843	2.215
The calyx	3.114	30.27	0	20.117
CSD	1.13	5.366	0.058	3.116
EPVL	2.358	11.3	0.005	6.337
Constant	−7.73	30.56	0	0.005
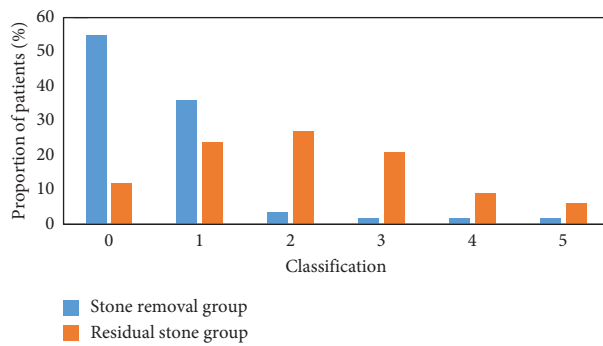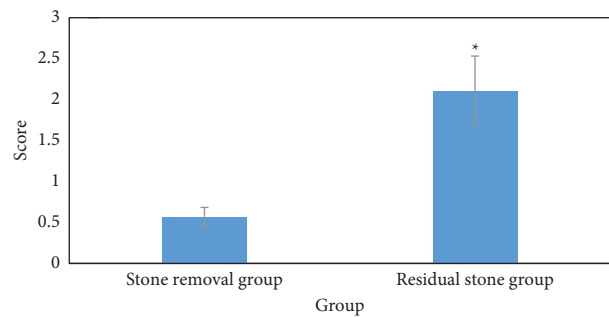
Figure 5: Comparison of SCR of RUSS stone scoring system. Note: (a) proportion of patients; (b) score results; compared with the stone clearance group, $^*P < 0.05$. "$*$" in "$^*P$" indicates that there is a significant difference between the two.

these treatments using well-known scoring methods. These methods are as follows.

(1) SSS: It is simple to assess and has a good predictive value, but it does not contain abnormal renal anatomy or spinal deformity and other factors that may affect the therapeutic effect of FURL, and its predictive ability has also to be validated with a prospective, multicenter study

(2) GSS grading: GSS is mainly graded according to the pelvicalyceal structure and stone distribution. For FURL, abnormal renal anatomy and different stone distribution will also affect the postoperative SCR of FURL. However, due to different positions, the surgical access channel adopted is inconsistent. GSS grading does not consider the factors such as stone size grading and stone cross-sectional area, which may be the reason for its poor effect in predicting the postoperative SCR of FURL

(3) RUSS score: The RUSS scoring system is simple and practical, but its indicators do not contain other indicators that may affect the SCR such as stone hardness and the degree of hydronephrosis

TABLE 4: Comparison of SCR of SSS.

	Stone clearance group ($n = 43$)	Stone residual group ($n = 25$)	Total	t	P
SSS	4.33 ± 1.13	8.26 ± 2.01	—	−4.324	0.002
		S		15.109	0.001
1	39 (90.7%)	15 (60.0%)	54 (79.4%)		
2	2 (4.7%)	8 (32.0%)	10 (14.7%)	#####	
3	2 (4.7%)	2 (8.0%)	4 (5.9%)		
		T		0.067	0.833
1	31 (72.1%)	14 (56.0%)	45 (66.2%)	#####	
2	12 (27.9%)	11 (44.0%)	23 (33.8%)		
		O		0.311	0.746
1	29 (67.4%)	15 (60.0%)	44 (64.7%)	#####	
2	14 (32.6%)	10 (40.0%)	24 (35.3%)		
		N		14.513	0.002
1	31 (72.1%)	13 (52.0%)	44 (64.7%)		
2	11 (25.6)	11 (44.0%)	22 (32.3%)	#####	
3	1 (2.3%)	1 (4.0%)	2 (2.9%)		
		E		5.538	0.041
1	34 (79.1%)	11 (44.0%)	45 (66.2%)	#####	
2	9 (20.9%)	14 (56.0%)	23 (33.8%)		

indicates that there is no relevant content.
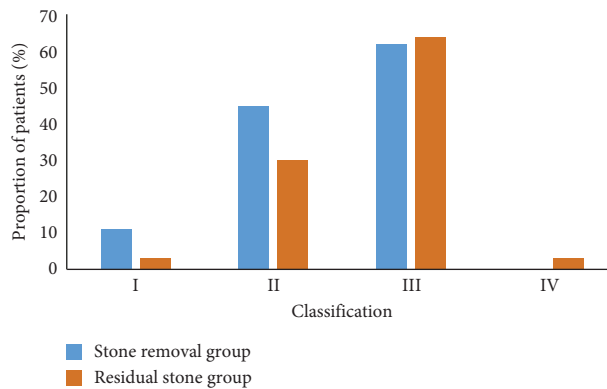
FIGURE 6: Comparison results of SCR of GSS.

(a)

(b)
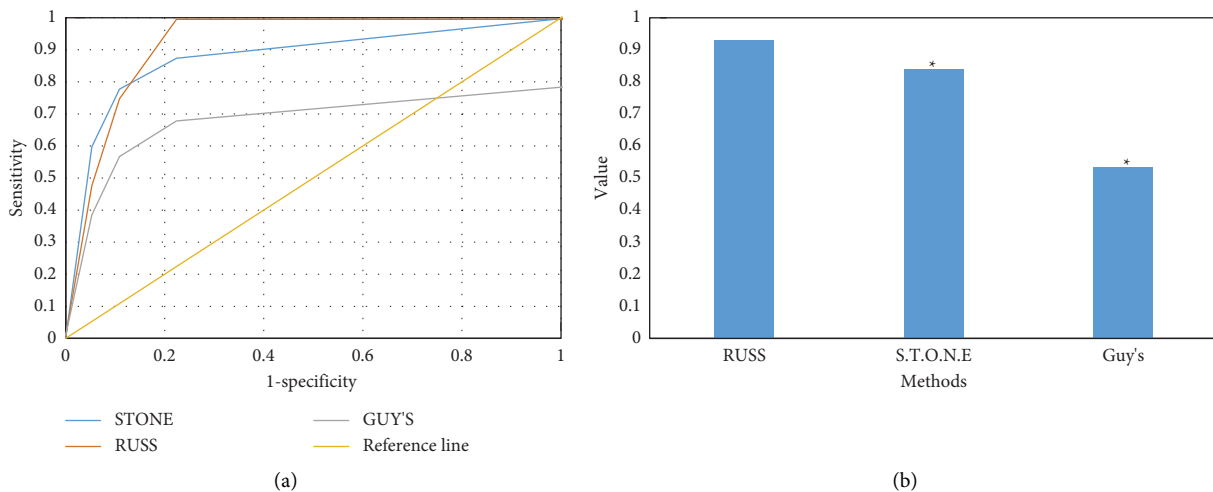
FIGURE 7: ROC curve analysis of FURL SCR by each scoring system. Note: (a) ROC curve; (b) area under the curve; compared with RUSS, *$P < 0.05$.

(4) CRS: The CRS nomogram can predict the renal stone complexity with great accuracy. Postoperative efficacy can be predicted accurately with CRS nephrolithometry

(5) S-ReSC: It was developed to predict the SFR after PCNL. It also describes the complexity of renal stones [17, 18].

In the previous researches, in order to compare the scoring, the score systems such as CRS, SSS, S-ReSC, and GSS for the prediction of SFS were compared. In our analysis, we have used the RUSS scoring system which has proven to be giving better results as compared to SSS and GSS. We also performed the RA and found that the stone location shows the strongest correlation of all other factors.

5. Conclusion

After the analysis, the results suggest that the factors such as location of stones (lower calyx), the CSD (20–30 mm), urinary tract infection, and EPVL are the main factors affecting the SCR after FURL. The stone location is found to have the strongest correlation of all. The RUSS scoring system has proven out to be the most accurate one for the evaluation of SCR after FURL, followed by the SSS and the GSS. Our current research compares the reliability of several existing scoring systems. In our future work, we will come up with the method for improving the existing scoring system. In the future, we will consider a bigger sample size for the analysis. We also will categorize the patients in different age groups, as kidney stone is being found in the younger population as well now a day.

Data Availability

The data can be made available on valid request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] L. F. Fontenelle and T. D. Sarti, "Kidney stones: treatment and prevention," *American Family Physician*, vol. 99, no. 8, pp. 490–496, 2019 Apr 15.

[2] O. Traxer and E. X. Keller, "Thulium fiber laser: the new player for kidney stone treatment? A comparison with Holmium: YAG laser," *World Journal of Urology*, vol. 38, no. 8, pp. 1883–1894, 2020 Aug.

[3] A. Carbone, Y. Al Salhi, A. Tasca et al., "Obesity and kidney stone disease: a systematic review," *Minerva Urologica e Nefrologica*, vol. 70, no. 4, pp. 393–400, 2018 Aug.

[4] P. Peerapen and V. Thongboonkerd, "Caffeine in kidney stone disease: risk or benefit?" *Advances in Nutrition*, vol. 9, no. 4, pp. 419–424, 2018 Jul 1.

[5] D. S. Goldfarb, "Empiric therapy for kidney stones," *Urolithiasis*, vol. 47, no. 1, pp. 107–113, 2019 Feb.

[6] M. Daudon, V. Frochot, D. Bazin, and P. Jungers, "Drug-induced kidney stones and crystalline nephropathy: pathophysiology, prevention and treatment," *Drugs*, vol. 78, no. 2, pp. 163–201, 2018 Feb.

[7] P. M. Ferraro, M. Bargagli, A. Trinchieri, and G. Gambaro, "Risk of kidney stones: influence of dietary factors, dietary patterns, and vegetarian-vegan diets," *Nutrients*, vol. 12, no. 3, p. 779, 2020 Mar 15.

[8] E. Letavernier and M. Daudon, "Vitamin D, hypercalciuria and kidney stones," *Nutrients*, vol. 10, no. 3, p. 366, 2018 Mar 17.

[9] C. Wang, L. Jin, X. Zhao, G. Li, and B. Xue, "Minimally invasive treatment of an ectopic kidney stone: a case report and literature review," *Journal of International Medical Research*, vol. 47, no. 9, pp. 4544–4550, 2019 Sep.

[10] C. D'Alessandro, P. M. Ferraro, C. Cianchi, M. Barsotti, G. Gambaro, and A. Cupisti, "Which diet for calcium stone patients: a real-world approach to preventive care," *Nutrients*, vol. 11, no. 5, p. 1182, 2019 May 27.

[11] L. Canat, H. A. Atalay, R. B. Degirmentepe et al., "Stone volume measuring methods: should the CT based three-dimensional-reconstructed algorithm be proposed as the gold standard? What did the three-dimensional printed models show us?" *Archivos Españoles de Urología*, vol. 72, no. 6, pp. 596–601, 2019.

[12] S. Yoshida, T. Hayashi, M. Morozumi, H. Osada, N. Honda, and T. Yamada, "Three-dimensional assessment of urinary stone on non-contrast helical computed tomography as the predictor of stonestreet formation after extracorporeal shock wave lithotripsy for stones smaller than 20 mm," *International Journal of Urology*, vol. 14, no. 7, pp. 665–667, 2007.

[13] R. P. Reimer, K. Klein, M. Rinneburger et al., "Manual kidney stone size measurements in computed tomography are most accurate using multiplanar image reformatations and bone window settings," *Scientific Reports*, vol. 11, no. 1, Article ID 16437, 2021.

[14] F. Richard, J. Marguin, J. Frontczak et al., "Evaluation and comparison of scoring systems for predicting stone-free status after flexible ureteroscopy for renal and ureteral stones," *PLoS One*, vol. 15, no. 8, Article ID e0237068, p. e0237068, 2020.

[15] M. Kaur, S. R. Sakhare, K. Wanjale, and F. Akter, "Early Stroke Prediction Methods for Prevention of Strokes," *Behavioural Neurology*, vol. 2022, Article ID 7725597, pp. 1–9, 2022.

[16] N. M. Mule, D. D. Patil, and M. Kaur, "A comprehensive survey on investigation techniques of exhaled breath (EB) for diagnosis of diseases in human body," *Informatics in Medicine Unlocked*, vol. 26, Article ID 100715, p. 100715, 2021.

[17] R. Ozbek, C. Senocak, H. B. Haberal, E. Damar, F. E. Sadioglu, and O. F. Bozkurt, "Comparison of scoring systems for predicting stone-free status and complications after retrograde ıntrarenal surgery," *World Journal of Urology*, vol. 39, no. 7, pp. 2741–2746, 2021.

[18] S. Azab and A. Osama, "Factors affecting lower calyceal stone clearance after Extracorporeal shock wave lithotripsy," *African Journal of Urology*, vol. 19, no. 1, pp. 13–17, 2013.

Research Article

Accuracy of Brain Computed Tomography Diagnosis by Emergency Medicine Physicians

Zohair Al Aseri [iD],[1,2] **Mohamed Al Aqeel,**[1] **Badr Aldawood,**[1] **Fahad Albadr,**[3]
Rawan Ghandour,[1] **Abdulaziz Al Mulaik,**[1] **Mohammed A. Malabarey,**[1] **and Anas Khan** [iD][1]

[1]*Department of Emergency Medicine, College of Medicine, King Saud University Medical City, King Saud University, Riyadh, Saudi Arabia*
[2]*Department of Critical Care, College of Medicine, King Saud University Medical City, King Saud University, Riyadh, Saudi Arabia*
[3]*Radiology and Medical Imaging, Neuroradiology, King Saud University Medical City, King Saud University, Riyadh, Saudi Arabia*

Correspondence should be addressed to Zohair Al Aseri; zalaseri@ksu.edu.sa

Objectives. The objective of this study is to prospectively analyze emergency physicians' (EP's) abilities to interpret noncontrast computed tomography (NCCT) brain images in a blinded fashion and assess whether they can make medical decisions solely based on their interpretations. *Methods.* A cross-sectional study was conducted at the emergency department (ED), King Saud University Medical City (KSU-MC), Saudi Arabia, over a period of one year, from May 2014 to May 2015. Any patient who underwent plain brain NCCT during the study period in our ED was included in this study. An independent attending neuroradiologist compared the EP's interpretations with the official final reports dictated by an on-call radiologist. *Results.* A brain NCCT prospective chart audit of 1,524 patients was interpreted by ED physicians (EP) at KSU-MC from 2014–2015. The ages of patients were between 14 and 107 years, and the mean ± SD age was 45.6 ± 22.1 years. Radiological brain lesions were confirmed by EPs and radiology physicians in 230 (15.09) and 239 (15.68) patients, respectively, out of which concordance was observed in 170 (71.13) cases, with a kappa value of $r = 0.675$. Normal, chronic, and nil acute reports were made by EPs and radiology physicians for 1,295 (84.97) patients and 1,285 (84.32) patients, respectively, out of which concordance was observed in 1,225 (95.33) cases, with a kappa value of $r = 0.672$. The study results demonstrated that the overall agreement between EPs and radiologist specialists was 91.6, with a kappa value of .675 ($p < 0.001$). *Conclusion.* Emergency physicians are moderately accurate at interpreting brain NCCT compared to radiologists. More research is needed to discover the most cost-effective technique for reducing the number of significant misinterpretations.

1. Introduction

Neurologic and traumatic complaints are frequently screened in the ED with the use of brain NCCT, which is required in both critical and noncritical cases [1]. EPs must respond promptly to trauma and other severe situations in accordance with the findings of related investigations since time is important [2]. Many EDs operate at a fast pace to avoid the formality of referrals and tracking radiological reports to save time and provide optimal management to their patients. However, some studies have implicated diagnostic imaging as a direct cause of increased ED length of stay [3, 4]. To improve clinical decision-making and treatment planning, the use of NCCT in patients with ED conditions might be beneficial via improved diagnostic confidence in the NCCT results [5]. Patients, healthcare providers, and managers might all benefit from an EP capable of appropriately understanding NCCT [6].

The discordance between a senior EP and a consultant radiologist was recorded at 14.8%, while the discordance

between a junior radiology trainee and a consultant radiologist was recorded at 40% [7]. Another study analyzed the interpretations of EPs and medical registrars and compared them to the final report of a radiologist, with a disagreement rate of 33% [8]. In a case series, the authors reached 13.4% disagreement between the interpretations of EP and two senior radiologists; however, there were no clinically relevant misinterpretations [9].

The current evidence shows inconsistent findings regarding the agreement between EPs and radiologists when it comes to brain NCCT. Therefore, in this study, we aimed to prospectively analyze EPs' abilities to interpret brain NCCT in a blinded fashion and assess whether they could make medical decisions solely based on their interpretations.

2. Methods

This cross-sectional study was conducted at the ED of King Saud University Medical City (KSU-MC). The study targeted 1,524 individuals who came to our ED while receiving brain NCCT and was conducted over a period of one year, from May 2014 to May 2015. Any patient who underwent plain brain NCCT during the study period in our ED was included in this study. We excluded only patients whose brain NCCT was interpreted by physicians other than EPs (on-call radiologists/admitting medical services, i.e., neurologists or neurosurgeons).

Ordering physicians are required to fill out patient identifying information, the study's initial interpretation, and the patient's final disposition plan before brain NCCT is performed. Only board-certified EPs with variable clinical experience were included in the evaluation. None of the EPs received further training or instructions for brain NCCT. After performing brain NCCT, the EP evaluated it and recorded their comments and dispositions on a form provided by a nurse. If the CT results were not documented on the EP report form when reported by the radiology expert, they were excluded from the study. Later, the forms were retrieved from the study's designated box. Finally, the radiological reports were tracked through the radiology system.

An independent attending neuroradiologist and two emergency medicine consultants compared the EP's interpretation with the official final reports dictated by the on-call radiologist. Upon reviewing each report, it was deemed either positive with no discrepancies or discrepant. Acute hemorrhage, acute/subacute infarction, evidence of space-occupying lesions, brain edema, evidence of cavernous sinus thrombosis, facial bone fracture, or acute hydrocephalus were all considered positive findings in our study since they might have altered the patient's immediate disposition. Positive NCCT with no discrepancies occurs when the ED physician and the on-call radiologist agree on the etiology. It is important to distinguish between fatal and nonfatal lesions because fatal ones need immediate attention and specialized consultation, while nonfatal ones do not. Data analysis was carried out using the Statistical Package for Social Science software (IBM SPSS Statistics Grad Pack 28.0). The sensitivity, specificity, concordance,

and kappa coefficient were calculated using the radiologist's judgment as the reference standard for evaluating the inter-rater reliability between the radiologist's report and EP's impression. Excellent agreement is defined as a kappa value of >0.75, with 0.40–0.75 being considered moderate, and <0.40 being regarded as poor. A p value of <0.05 was considered significant.

3. Results

As shown in Table 1, an emergency brain NCCT prospective chart audit of 1,524 patients was interpreted by EPs at KSU-MC from 2014-2015. The patients were over 14 years and presented to the ED undergoing NCCT brain as per EP discretion. The ages of the patients were between 14 and 107 years, and the mean ± SD age was 45.6 ± 22.1 years.

The chief complaints in the patients were trauma (27.4%), loss of consciousness (23.1%), headache (22.6%), weakness/numbness (16.8%), dizziness (9.4%), seizure (8.7%), nausea and vomiting (5.9%), dysphasia, aphasia, and speech difficulty (5.0%). Visual disturbances, difficulty walking/ataxia, craniopathies, vertigo, amnesia, insomnia, and urine incontinence were recorded at 18.9%.

The indications for ordering brain NCCT were to look for intracranial bleeding (877, 57.5%), ischemic stroke (435, 28.5%), and space-occupying lesion/metastasis (197, 12.9%), while 27 (1.9%) cases were for hydrocephalus, cavernous sinus thrombosis, facial bone fracture, edema, and shift. Other 408 (26.8%). 132 (8.7%) cases had not been documented.

The ED observed 1–3 findings per patient, thereby documenting 256 lesions in 230 (15.1%) patients, while most reports (1,295, 84.9%) were considered normal, chronic, or nil acute. Similarly, radiologists had observed 1–3 findings per patient, thereby documenting 271 lesions in 239 (15.7%) patients, while most of the reports (1,285, 84.3%) were deemed normal.

As shown in Table 2, brain NCCT findings were classified into ten findings. The EP's report was tallied with the radiologist's confirmed report for each of the specified findings. Inter-rater agreement was represented by the percentage sensitivity score, and the measure of association between EPs and radiologists was agreed upon with consideration of the kappa test results.

EPs and radiologists reported intracranial bleeding in 24 (1.57) and 26 (1.71) patients, respectively, among which concordance was observed in 18 (69.23) cases, with a kappa value of $r = 0.715$. Subdural hemorrhage was reported in 20 (1.31) and 18 (1.18) patients, respectively, among which concordance was observed in 12 (66.67) cases, with a kappa value of $r = 0.627$. Subarachnoid hemorrhage (SAH) was observed in 12 (0.79) and 7 (0.46) patients, respectively, among which concordance was observed in six (85.71) cases, with a kappa value of $r = 0.629$. Ischemic stroke was reported in 120 (7.87) and 123 (8.07) patients, respectively, among which concordance was observed in 74 (60.16) cases, with a kappa value of $r = 0.575$. Space-occupying lesions were reported in 22 (1.44) and 30 (1.97) patients, respectively, among which concordance was observed in 18

TABLE 1: Descriptive statistics.

Characteristics	Description	N (%)
Age (year)	Mean ± SD	45.6 ± 22.1
	Minimum-maximum	14–107
Chief complaints	Trauma	417 (27.4)
	Loss of consciousness	352 (23.1)
	Headache	344 (22.6)
	Weakness/numbness	256 (16.8)
	Dizziness	143 (9.4)
	Seizure	133 (8.7)
	Nausea and vomiting	90 (5.9)
	Dysphasia, aphasia, speech difficulty	76 (5.0)
	Visual disturbances	45 (3.0)
	Difficulty walking/ataxia	36 (2.4)
	Craniopathies	23 (1.5)
	Vertigo	21 (1.4)
	Amnesia	10 (0.7)
	Insomnia	7 (0.5)
	Urine incontinence	3 (0.2)
	Others	140 (9.2)
Indications	Intracranial bleeding (SAH, SDH, hemorrhagic stroke, etc.)	877 (57.5)
	Ischemic stroke	435 (28.5)
	Space-occupying lesion/metastasis	197 (12.9)
	Hydrocephalus	13 (0.9)
	Cavernous sinus thrombosis	6 (.4)
	Facial bone fracture	4 (0.3)
	Edema, shift	4 (0.3)
	Others	408 (26.8)
	Not documented	132 (8.7)
EP findings	Nil	1295 (84.9)
	One	206 (13.5)
	Two	22 (1.4)
	Three	2 (0.1)
	Total findings	230
Radiology findings	Nil	1285 (84.3)
	One	211 (13.8)
	Two	24 (1.6)
	Three	4 (0.3)
	Total findings	239
Concordance (including nil)		**1395 (91.6)**

(60.00) cases, with a kappa value of $r = 0.687$. Edema and shift were reported in 24 (1.44) and 20 (1.97) patients, respectively, among which concordance was observed in 14 (70.00) cases, with a kappa value of $r = 0.631$. Hydrocephalus was reported in six (0.39) and seven (0.46) patients, respectively, among which concordance was observed in three (42.86) cases, with a kappa value of $r = 0.459$. Cavernous sinuses thrombosis was reported in four (0.26) and three (0.20) patients, respectively, among which concordance was observed in three (85.71) cases, with a kappa value of $r = 0.857$. Facial bone fractures were reported in 18 (1.18) and 32 (2.10) patients, respectively, among which concordance was observed in 17 (53.13) cases, with a kappa value of $r = 0.675.10$-epidural hematoma was reported in six (0.39) and five (0.33) patients, respectively, among which concordance was observed in five (90.19) cases, with a kappa value of $r = 0.909$.

There were normal, chronic, and nil acute ED reports for 1,295 (84.97) patients and radiologist reports for 1,285 (84.32) patients, respectively, out of which concordance was

observed in 1,225 (95.33) cases, with a kappa value of $r = 0.672$.

CT brain lesions were confirmed by EPs in 230 (15.09) patients and radiologists in 239 (15.68) patients, out of which concordance was observed in 170 (71.13) cases, with a kappa value of $r = 0.675$. The overall agreement between EPs and radiologist specialists was 91.6, with a kappa value of .675 ($p < 0.001$).

4. Discussion

For traumatic patients, accurate interpretation of NCCT brain abnormalities is critical for prompt and appropriate management. Previously, many studies have been conducted to evaluate the accuracy of EPs' interoperations with NCCT and other imaging modalities; however, the design and methodology of these studies vary considerably [10, 11]. The interpretation of plain X-rays was the main focus of most of these studies rather than CT, with an overall discordance rate ranging from 0.95% to 16.8% [12, 13]. It was reported that the total discordance rate

TABLE 2: Inter-rater agreement between the emergency physician and radiologist.

Characteristics (N = 1524)	Cell frequency						Measures of association					Agreement
Site	ED findings	RAD findings	ED+/RAD+	ED+/RAD−	ED−/RAD+	ED−/RAD−	Sensitivity	Specificity	PPV	NPV	Accuracy	Kappa
Intracranial bleeding	24	26	18	6	8	1492	69.23	99.60	75.00	99.50	99.10	0.715
Subdural hemorrhage	20	18	12	8	6	1498	66.67	99.47	60.00	99.60	99.10	0.627
Subarachnoid hemorrhage	12	7	6	6	1	1511	85.71	99.60	50.00	99.90	99.50	0.629
Ischemic stroke	120	123	74	46	49	1355	60.16	96.72	61.67	96.50	93.80	0.575
Space-occupying lesions	22	30	18	4	12	1490	60.00	99.73	81.82	99.20	99.00	0.687
Edema, shift	24	20	14	10	6	1494	70.00	99.34	58.33	99.60	99.00	0.631
Hydrocephalus	6	7	3	3	4	1514	42.86	99.80	50.00	99.70	99.50	0.459
Cavernous sinus thrombosis	4	3	3	1	1	1520	85.71	99.93	75.00	100.00	99.90	0.857
Facial bone fracture	18	32	17	1	15	1491	53.13	99.93	94.44	99.00	99.00	0.675
Epidural	6	5	5	1	1	1518	90.91	99.93	83.33	100.00	99.90	0.909
Normal, chronic, nil acute	1295	1285	1225	70	60	169	95.33	70.71	94.59	73.80	91.50	0.672
Overall confirmed cases	230	239	170	60	69	1225	71.13	95.33	73.91	94.70	91.50	0.675

between the interpretations of NCCT by EPs was 37.1% in the only comparable prospective study [14]. In this study, our findings showed 91.6% concordance between EPs and radiologists. The agreement between EP and radiologist specialists was good, with a kappa value of .675 ($p < 0.001$).

Similar to our findings, Khan et al. showed that the concordance between EPs and radiologists was observed in 87.14% of the interpreted NCCT brain images with good agreement (kappa = 0.64) and a high degree of accuracy of 90.5%. They also showed that the false-negative rate of EPs was 3.6% [17]. In an English retrospective study, Mucci et al. investigated the accuracy of five EPs in the interpretation of 100 NCCT brain images. Their findings demonstrated that the overall agreement was 86.6%, and the false-negative rate was 4.2% [10]. It is common to see 1%–3% false-negative rates in published studies; however, in some studies, they may reach 11%. Using a set of completely abnormal images, Vincent et al. reported that 35% of EPs made mistakes [19]. Nevertheless, clinical outcomes are rarely affected by missed diagnoses during the time between initial interpretation and radiological evaluation [17–19].

In an Australian study, the authors demonstrated that EPs were able to accurately interpret 85.20% of cases, with a kappa value of 0.69. Out of discordant events, 41% of cases appeared with potential or definitive complications. Among the scans that the radiologist deemed abnormal, the discordance rate was higher. They also highlighted that there was no significant difference in NCCT interpretation accuracy based on the EP's degree of experience or training [7]. The likeliest reason is that CT interpretation accuracy is mostly determined by formal training rather than emergency medical skills. Seniority does not seem to have any effect on physicians' ability to report either plain films or CT scans [19, 20]. With only one to two hours of training, EPs

have shown a statistically significant increase in NCCT reporting accuracy [21]. In addition, the first year of radiology training has a large impact on trainees' accuracy, but this influence diminishes with time as variations between individuals become more apparent [22].

According to Khoo and Duffy, only around two-thirds of senior EPs' interpretations of NCCT scans are accurate. The proportion of "abnormal" scans in their study population was 26%, which yielded a decent negative predictive value of 82.3% [9]. How much precision is needed to ensure safe practice remains unclear. Even though their clinically significant misinterpretation rate was just 6.1%, Arendts et al. judged that their level of accuracy was "no better than moderately good" [8]. Alfaro et al., on the other hand, reported that despite a very high incidence of misinterpretation, just 0.6% of patients were treated improperly, and none had an unfavorable result in the chart review [16].

SAH is a life-threatening condition that commonly results in brain damage and even death [23–25]. It may be difficult to diagnose and could present with nonspecific complaints [26]. There is a greater risk of re-bleeding and associated death if SAH is diagnosed and treated late [27, 28]. Our study showed an accuracy of 99.50 for diagnosing SAH. Some studies have shown that the misdiagnosis of SAH varies between 12% and 51% [29–31]. Access to diagnostic resources, physician experience, and patient acuity are all important risk factors for misdiagnosis. A large cohort study showed that the rate of misdiagnosis of cerebral cavernous sinuses thrombosis was 3.6%, which was associated with an increased risk of a longer length of hospital stay, an unfavorable discharge disposition, intracerebral hemorrhage, and in-hospital mortality [32]. In our study, the accuracy of diagnosing cavernous sinus thrombosis was 99.90, but there were only four cases. A cross-sectional study

of 2,288 cases showed that fractures were the most common conditions associated with diagnostic errors in ED (44%), followed by intracranial bleeding (6%). They also stated that human mistakes, healthcare professionals' inadequate skills or knowledge, and inability to comply with protocols were among the major factors for these errors [33, 34]. Our study showed an accuracy of 99% in diagnosing intracranial bleeding and facial bone fractures.

We acknowledge that our study has some limitations, including the single-center setting, which may induce the risk of selection bias. Therefore, multicenter studies are required to investigate the predictors of EPs' interpretations, including the variation between healthcare systems, access to diagnostic tools, compliance of physicians with the protocol, and training received [35–37].

In conclusion, our EPs were moderately accurate in interpreting brain NCCT compared to radiologists. More research is needed to discover the most cost-effective technique for reducing the number of significant misinterpretations. Brain NCCT interpretation instruction sessions may significantly enhance EPs' accuracy.

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] H. Shah, B. Jarwani, and M. Gjjar, "Can ed physician interpret NCCT brain reliably in head injury victims's emergency management?" *Journal of Evolution of Medical and Dental Sciences*, vol. 4, no. 8, pp. 1306–1311, 2015.

[2] T. L. Rodziewicz and B. H. J. Houseman, *Medical Error Reduction and Prevention*, StatPearls, Tampa, FL, USA, 2022.

[3] M. Goloback, D. M. McCarthy, M. Schmidt, J. G. Adams, and P. S. Pang, "ED operational factors associated with patient satisfaction," *The American Journal of Emergency Medicine*, vol. 33, no. 1, pp. 111-112, 2015.

[4] A. Ryan, K. Hunter, K. Cunningham et al., "STEPS: lean thinking, theory of constraints and identifying bottlenecks in an emergency department," *Irish Medical Journal*, vol. 106, no. 4, pp. 105–107, 2013.

[5] B. K. Menon and A. M. Demchuk, "Computed tomography angiography in the assessment of patients with stroke/TIA," *The Neurohospitalist*, vol. 1, no. 4, pp. 187–199, 2011.

[6] M. Badr, S. Al-Otaibi, N. Alturki, and T. Abir, "Detection of heart arrhythmia on electrocardiogram using artificial neural networks," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1094830, 10 pages, 2022.

[7] Institute of Medicine (US) Committee on the Health Professions Education Summit, "The core competencies needed for health care professionals," in *Health Professions Education: A Bridge to Quality*, A. C. K. E. Greiner, Ed., National Academies Press (US), Washington, DC, USA, 2003.

[8] G. Arendts, A. Manovel, and A. Chai, "Cranial CT interpretation by senior emergency department staff," *Australasian Radiology*, vol. 47, no. 4, pp. 368–374, 2003.

[9] N. C. Khoo and M. Duffy, "Out of hours non-contrast head CT scan interpretation by senior emergency department medical staff," *Emergency Medicine Australasia*, vol. 19, no. 2, pp. 122–128, 2007.

[10] B. Mucci, C. Brett, L. S. Huntley, and M. K. Greene, "Cranial computed tomography in trauma: the accuracy of interpretation by staff in the emergency department," *Emergency Medicine Journal*, vol. 22, no. 8, pp. 538–540, 2005.

[11] E. J. Klein, M. Koenig, D. S. Diekema, and W. Winters, "Discordant radiograph interpretation between emergency physicians and radiologists in a pediatric emergency department," *Pediatric Emergency Care*, vol. 15, no. 4, pp. 245–248, 1999.

[12] M. Maray, M. Alghamdi, and M. B. Alazzam, "Diagnosing cancer using IOT and machine learning methods," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9896490, 9 pages, 2022.

[13] L. A. Nitowski, R. E. O'Connor, and C. L. Reese, "The rate of clinically significant plain radiograph misinterpretation by faculty in an emergency medicine residency program," *Academic Emergency Medicine*, vol. 3, no. 8, pp. 782–789, 1996.

[14] P. G. Herman and S. J. Hessel, "Accuracy and its relationship to experience in the interpretation of chest radiographs," *Investigative Radiology*, vol. 10, no. 1, pp. 62–67, 1975.

[15] T. M. Nolan, F. Oberklaid, and D. Boldt, "Radiological services in a hospital emergency department—an evaluation of service delivery and radiograph interpretation," *Journal of Paediatrics and Child Health*, vol. 20, no. 2, pp. 109–112, 1984.

[16] D. Alfaro, M. A. Levitt, D. K. English, V. Williams, and R. Eisenberg, "Accuracy of interpretation of cranial computed tomography scans in an emergency medicine residency program," *Annals of Emergency Medicine*, vol. 25, no. 2, pp. 169–174, 1995.

[17] A. Khan, S. Qashqari, and A.-A. Al-Ali, "Accuracy of non-contrast CT brain interpretation by emergency physicians: a cohort study," *Pakistan Journal of Medical Sciences*, vol. 29, no. 2, pp. 549–553, 2013.

[18] N. A. Qader Osman, S. H. Al-Ziyadi, M. B. Alazzam, S. Z. Alshawwa, and M. A. Rahman, "Machine learning of ZnO interaction with immunoglobulins and blood proteins in medicine," *Journal of Healthcare Engineering*, vol. 2022, Article ID 4062974, 6 pages, 2022.

[19] C. A. Vincent, P. A. Driscoll, R. J. Audley, and D. S. Grant, "Accuracy of detection of radiographic abnormalities by junior doctors," *Emergency Medicine Journal*, vol. 5, no. 2, pp. 101–109, 1988.

[20] J. Wardrope and P. M. Chennells, "Should all casualty radiographs be reviewed?" *British Medical Journal Publishing Group*, vol. 290, no. 6482, pp. 1638–1640, 1985.

[21] S. R. Eachempati, N. Flomenbaum, C. Seifert, E. Fischer, L. J. Hydo, and P. S. Barie, "Alterations of preliminary readings on radiographic examinations minimally affect outcomes of trauma patients discharged from the emergency department," *The Journal of Trauma, Injury, Infection, and Critical Care*, vol. 48, no. 4, pp. 654–658, 2000.

[22] F. E. Mayhue, D. D. Rust, J. C. Aldag, A. M. Jenkins, and J. C. Ruthman, "Accuracy of interpretations of emergency department radiographs: effect of confidence levels," *Annals of Emergency Medicine*, vol. 18, no. 8, pp. 826–830, 1989.

[23] M. C. Gratton, J. A. Salomone, and W. A. Watson, "Clinically significant radiograph misinterpretations at an emergency medicine residency program," *Annals of Emergency Medicine*, vol. 19, no. 5, pp. 497–502, 1990.

[24] M. A. Levitt, R. Dawkins, V. Williams, and S. Bullock, "Abbreviated educational session improves cranial computed tomography scan interpretations by emergency physicians," *Annals of Emergency Medicine*, vol. 30, no. 5, pp. 616–621, 1997.

[25] J. A. Espinosa and T. W. Nolan, "Reducing errors made by emergency physicians in interpreting radiographs: longitudinal study," *BMJ Publishing Group*, vol. 320, no. 7237, pp. 737–740, 2000.

[26] Y. Roos, R. J. De Haan, L. F. M. Beenen, R. J. M. Groen, K. W. Albrecht, and M. Vermeulen, "Complications and outcome in patients with aneurysmal subarachnoid haemorrhage: a prospective hospital based cohort study in The Netherlands," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 68, no. 3, pp. 337–341, 2000.

[27] A. Hijdra, R. Braakman, J. Van Gijn, M. Vermeulen, and H. Van Crevel, "Aneurysmal subarachnoid hemorrhage. Complications and outcome in a hospital population," *Stroke*, vol. 18, no. 6, pp. 1061–1067, 1987.

[28] N. J. Solenski, E. C. J. Haley, N. F. Kassell et al., "Medical complications of aneurysmal subarachnoid hemorrhage: a report of the multicenter, cooperative aneurysm study," *Critical Care Medicine*, vol. 23, no. 6, pp. 1007–1017, 1995.

[29] M. B. Alazzam, A. T. Al-Radaideh, R. A. Alhamarnah, F. Alassery, F. Hajjej, and A. Halasa, "A survey research on the willingness of gynecologists to employ mobile health applications," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1220374, 7 pages, 2021.

[30] M. J. Vermeulen and M. J. Schull, "Missed diagnosis of subarachnoid hemorrhage in the emergency department," *Stroke*, vol. 38, no. 4, pp. 1216–1221, 2007.

[31] A. S. Lord, L. Fernandez, J. M. Schmidt et al., "Effect of rebleeding on the course and incidence of vasospasm after subarachnoid hemorrhage," *Neurology*, vol. 78, no. 1, pp. 31–37, 2012.

[32] C. Tang, T.-S. Zhang, and L.-F. Zhou, "Risk factors for rebleeding of aneurysmal subarachnoid hemorrhage: a meta-analysis," *PLoS One*, vol. 9, no. 6, Article ID e99536, 2014.

[33] P. Vannemreddy, A. Nanda, R. Kelley, and M. K. Baskaya, "Delayed diagnosis of intracranial aneurysms: confounding factors in clinical presentation and the influence of misdiagnosis on outcome," *Southern Medical Journal*, vol. 94, no. 11, pp. 1108–1111, 2001.

[34] R. G. Kowalski, J. Claassen, K. T. Kreiter et al., "Initial misdiagnosis and outcome after subarachnoid hemorrhage," *JAMA. United States*, vol. 291, no. 7, pp. 866–869, 2004.

[35] P. L. Mayer, I. A. Awad, R. Todor et al., "Misdiagnosis of symptomatic cerebral aneurysm. Prevalence and correlation with outcome at four institutions," *Stroke*, vol. 27, no. 9, pp. 1558–1563, 1996.

[36] A. L. Liberman, G. Gialdini, E. Bakradze, A. Chatterjee, H. Kamel, and A. E. Merkler, "Misdiagnosis of cerebral vein thrombosis in the emergency department," *Stroke*, vol. 49, no. 6, pp. 1504–1506, 2018.

[37] F. Hussain, A. Cooper, A. Carson-Stevens et al., "Diagnostic error in the emergency department: learning from national patient safety incident report analysis," *BMC Emergency Medicine*, vol. 19, no. 1, p. 77, 2019.

Hindawi

Review Article

A Systematic Review and Meta-Analysis of the Efficacy of Uterine Artery Embolization in the Treatment of Endometriosis

Li Ma [ID],[1] **Bingxin Wen,**[2] **and Zhenhua Wen** [ID][1]

[1]*Operating Room,the First Hospital of China Medical University, Shenyang 110001, Liaoning, China*
[2]*Department of General Surgery,Jin Qiu Hospital of Liaoning, Shenyang 110016, Liaoning, China*

Correspondence should be addressed to Zhenhua Wen; wzhenhua75700@163.com

Objective. To compare the efficacy of uterine artery embolization (UAE) with traditional methods for treating endometriosis. *Methods.* The randomized controlled trials of uterine artery embolization and other medical treatments for endometriosis in PubMed, Embase, Web of Science, Cochrane Library, China Journal Full-Text Database, Wanfang Database, VIP Database, and China Biomedical Literature Database were retrieved by computer. The search time was up to June 2022. The quality of articles was evaluated by Cochrane ROB 2.0, and meta-analysis was performed by Stata15.1 software. *Results.* 7 studies were finally included. Meta-analysis showed that the serum CA125 level after uterine artery embolization was significantly lower than that in the control group (SMD = −0.85, 95%CI (−1.12, −0.59)], and the postoperative visual analogue scale (VAS) of dysmenorrhea was significantly lower than that in the control group (SMD = −1.86, 95%CI (−2.21, −1.50)) There was no significant difference in the effective rate, FSH level, E2 level, and LH level between the two groups. *Conclusion.* UAE can effectively reduce the VAS score of dysmenorrhea and serum CA125 level for treating endometriosis. However, due to the limitation of the quality of included articles, more large sample size and high quality RCTs are needed to provide stronger evidence-based medicine evidence for clinical practice.

1. Foreword

Endometriosis (EMs) refers to inflammatory diseases in which endometrial tissue with growth function appears in the lining endometrium of the uterine cavity as well as in parts other than the uterus, affecting about 5% to 10% of women worldwide and women of reproductive age [1], and is regarded as a modern disease [2]. The main clinical manifestations are pain, mass, infertility, and symptoms of pelvic pain occurring in 70%–80% of patients [3]. Long-term management is needed to ease pain [4, 5]. At present, western medicine mainly adopts hormone therapy alone or combined surgery. Although it can slow down the development of endometriosis lesions to a certain extent, achieve the purpose of relieving the disease and inhibiting recurrence, the damage of ovarian function by surgery, the damage to the pelvic environment, as well as the unavoidable high recurrence rate after surgery and the

untoward reactions existing in the process of hormone drug therapy such as irregular vaginal bleeding, menopausal symptoms, and thrombotic risk all affect its clinical efficacy and long-term clinical use. In view of its unclear pathogenesis, the efficacy of surgical treatment alone is not satisfactory, there has been a long lack of more effective treatment. In recent years, minimally invasive therapy represented by uterine artery embolization has gradually emerged. UAE has been first applied in the diagnosis and treatment of female surgical pelvic bleeding and gynecological diseases, which shows the advantages of less trauma, rapid postoperative recovery, and preservation of the uterus and exact effect and conforms to the people-oriented medical ethical concept. In this study, we systematically evaluated randomized controlled trials of UAE compared with other treatments for endometriosis in order to provide evidence-based evidence for clinicians to select reasonable surgical methods in the future.

2. Materials and Methods

2.1. Literature Retrieval. Databases in English were searched: PubMed, Embase, Web of Science, and Cochrane Library, while databases in Chinese were searched: China Journal Full-text Database, Wanfang Database, VIP Database, and China Biomedical Literature Database. The RCTs related to EMs after uterine artery embolization were retrieved from the established until June 2022. Use subject to merge search methods for free words. The search terms were *[uterine artery] [embolization endometriosis] [chocolate cyst] [Randomized].*

2.2. Inclusion Criteria

- (i) RCTs of UAE combined with other therapies for the treatment of Ems published in Chinese or English
- (ii) Patients diagnosed with endometriosis according to diagnostic criteria
- (iii) uterine artery embolization was adopted in the observation group, while other treatments were adopted in the control group

2.3. Exclusion Criteria

- (i) Study on the treatment inconsistent with the research purpose between the observation group and the control group
- (ii) Articles that are published repeatedly or with incorrect data
- (iii) Articles for which outcome indicators could not be extracted

2.4. Risk Assessment of Literature Bias. The Cochrane ROB 2.0 [6] was used to assess the risk of publication bias from the aspects of randomization process, deviation from established interventions, lack of outcome data, measurement of results, selective reporting bias, and total bias. It was conducted independently and cross-checked by two researchers.

2.5. Data Extraction. Literature screening was conducted by two colleagues independently, and the data were extracted and cross-checked. When differences of opinion arose, the two researchers discussed and resolved the differences with the third researcher. The extracted contents mainly include the following:

- (i) Basic information of the literature: author, year of publication, title, and source of the article
- (ii) The evaluation factors of bias risk were as follows: random method and blind method
- (iii) Intervention measure
- (iv) Outcome indicators

2.6. Statistical Analysis. Stata 15.1 software was used to process and analyze the included studies, enumeration data

were expressed by relative risk (*RR*), measurement data were expressed by standardized mean difference (*SMD*), and 95% CI was used to express the results. Heterogeneity was quantitatively evaluated using χ^2 test and I^2. If there was no significant heterogeneity ($P > 0.1$, $I^2 < 50\%$), fixed-effect model was used for meta-analysis, while sensitivity analysis or subgroup analysis was used to investigate the source of heterogeneity. Random effect model was used for analysis when significant clinical heterogeneity was excluded. $P < 0.05$ was considered statistically significant. Egger's test was used to quantitatively detect publication bias.

3. Result

3.1. Literature Retrieval Results. A total of 338 articles were retrieved. After deduplication and preliminary screening, the remaining 230 articles were first screened, and finally, 7 articles were included. As shown in Figure 1, the basic situation of the inclusion is shown in Table 1.

3.2. Risk of Inclusion Study Bias. 7 articles with a total of 524 patients were included. Outcome indicators: total efficiency was reported in 5 articles, serum CA125 levels were measured in 3 articles, dysmenorrhea visual analogue scale was reported in 2 articles, and sex hormone levels were measured in 2 articles. 2 articles had a low risk of the randomization process, 2 articles had a high risk of the randomization process, and 3 articles only mentioned randomization, none mentioned deviation from established interventions, none lacked outcome data, had a low risk of outcome measurement, and had a low risk of selective reporting bias as shown in Figures 2 and 3.

3.3. Total Effective Rate. According to literature [14], it is considered to be effective when the clinical symptoms of the patient relieve or disappear, the pelvic mass shrinks or does not increase, and the clinical symptoms of the patient do not worsen after 3 months of drug discontinuation. Patient's clinical symptoms and signs were improved, or even worsened, and were considered invalid.

364 patients were included in the 5 studies to compare the difference in the total effective rate between the two groups, with statistically significant heterogeneity between the literature stuides: $I^2 = 0.0\%$, $P = 0.995$. Therefore, the fixed-effect model was used for analysis. The results showed that there was no significant difference in the effective rate of treatment of endometriosis with UAE and other treatment methods: (RR = 1.13, 95%CI (0.96, 1.34), $P < 0.05$), as shown in Figure 4.

3.4. Serum CA125 Level. 244 patients were included in the 3 RCTs, with statistically significant heterogeneity between the literature stuides: $I^2 = 18.7\%$, $P = 0.292$. Therefore, the fixed-effect model was used. Meta-analysis showed that the serum CA125 level in the UAE group was lower than that in the control group (SMD = −0.85, 95%CI (−1.12, −0.59)),
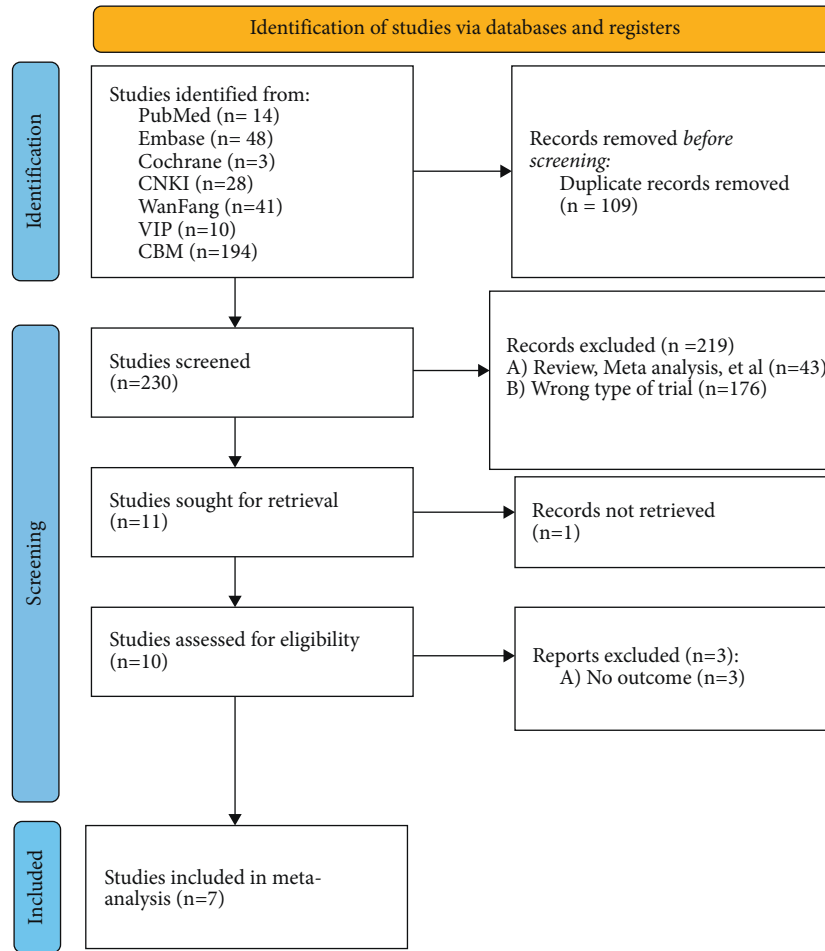
FIGURE 1: Screening plot.

TABLE 1: Baseline information of the literature.

Study	Year	Age		Total		Interventions		Outcome measurements
		T	C	T	C	T	C	
Liang CP et al. [7]	2015	36.0 ± 5.0	34.0 ± 6.0	40	40	UAE	Traditional surgery	②⑤
Lu CJ et al. [8]	2015	32.2 ± 6.1		30	30	UAE	Lesion resection	①
Chen MG et al. [9]	2019	35.55 ± 3.91	36.21 ± 4.16	30	30	UAE	Traditional surgery	①②⑤
Chen FH et al. [10]	2018	37.57 ± 8.12	37.64 ± 8.20	50	48	UAE	Ibuprofen	①③
Wang YL et al. [11]	2015	32.16 ± 6.71		42	42	UAE	Ultrasound ablation	①②⑤
Hu B et al. [12]	2021	32.4 ± 11.2	31.6 ± 10.8	31	31	UAE	Traditional surgery	①②
Zhang XY et al. [13]	2021	35.85 ± 5.06	36.06 ± 5.18	43	37	UAE	Lesion resection	①②③⑤

UAE: uterine artery embolization; T: trial; C: control; ① total efficiency; ② serum CA125 level; ③ dysmenorrhea VAS score; ④ sex hormone levels (FSH, LH, and E2); ⑤ others.

$P < 0.05$, indicating that the difference was statistically significant, as shown in Figure 5.

3.5. VAS Score of Dysmenorrhea. 178 patients were included in the 2 RCTs. The heterogeneity test results showed that the components had relatively small heterogeneity: $I^2 = 0.0\%$, $P = 0.416$. Therefore, the fixed-effect model was adopted. Meta-analysis showed that the VAS score of dysmenorrhea in the UAE group was lower than that in the control group

(SMD $= -1.86$, 95%CI $(-2.21, -1.50)$), $P < 0.05$, indicating a statistically significant difference, as shown in Figure 6.

3.6. Sex Hormone Levels. 160 patients were included in 2 literature studies. E2 level was highly heterogeneous among studies: $I^2 = 96.8\%$. Random effect model was used for analysis. Meta-analysis showed that there was no significant difference in E2 level between UAE and other treatment methods for endometriosis (SMD $= -1.32$, 95%CI $(-3.33$,
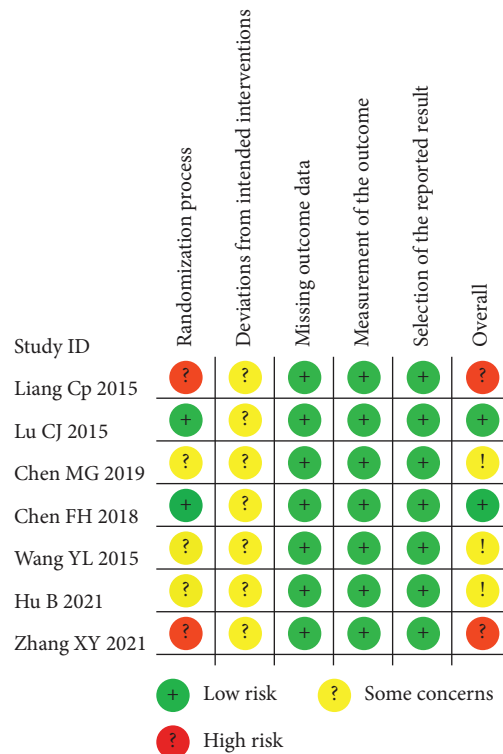
FIGURE 2: Detailed plot of bias of 7 included literature studies: based on Cochrane ROB 2.0.
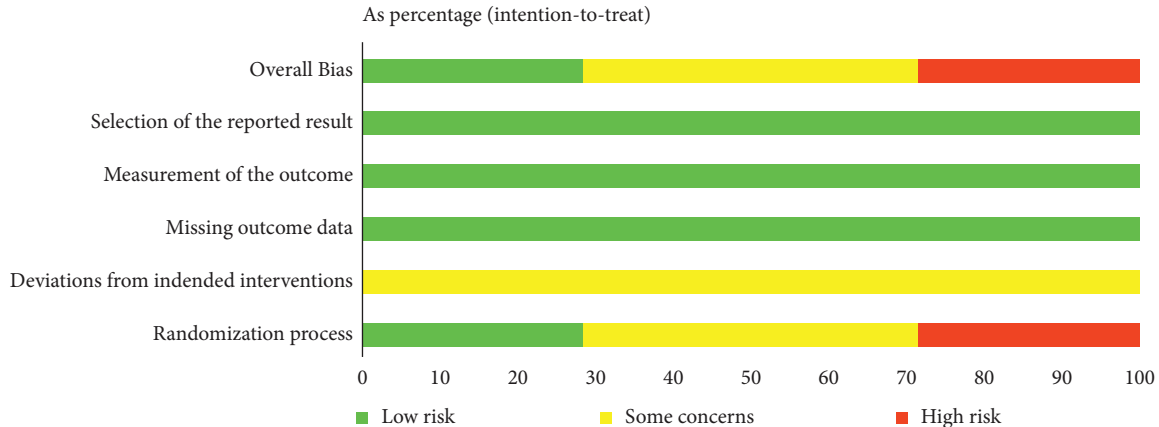
FIGURE 3: Summary plot of bias for the 7 included literature studies: based on Cochrane ROB 2.0.

0.69), $P = 0.197$). The results of FSH level (SMD = −0.23, 95%CI (−0.54, 0.09)) and LH level (SMD = 0.39, 95%CI (−0.50, 1.29)) showed no significant difference.

3.7. Publication Bias. The total effective rate was tested for potential publication bias by Egger's Test. The results of Egger's Test analysis are shown in Figure 7, $P = 0.752$, indicating that there was no significant publication bias.

4. Discussion

Endometriosis is a common disease and multiple diseases in women of childbearing age, and its pathogenesis is dominated by Sampson menstrual blood reflux implantation. Endometrium reflux to the pelvic cavity requires adhesion, invasion, and vascular formation to obtain implantation growth and finally develop lesions; traits of eutopic endometrium play a decisive role, that is "eutopic endometrium determinism" [15]; other pathogeneses include metaplasia of coelomic epithelium, immune-inflammation, abnormal expression of estrogen and progesterone receptors, and genetic factors [16, 17]. Endometriosis has familial aggregation, and women with endometriosis in first-degree relatives have a relatively 7–10-fold increased risk of endometriosis. Clinically, it is usually treated with surgery or combined hormonal drugs, such as the use of nonsteroidal
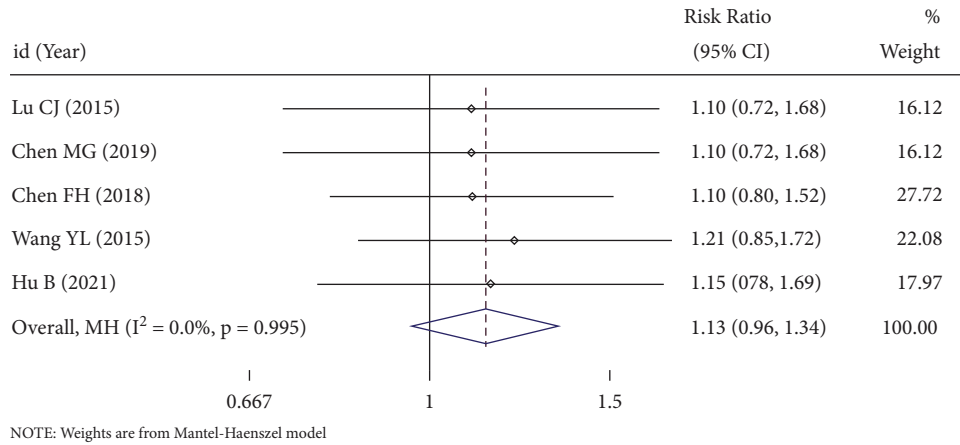
FIGURE 4: Forest plot of combined results for total effective rate.
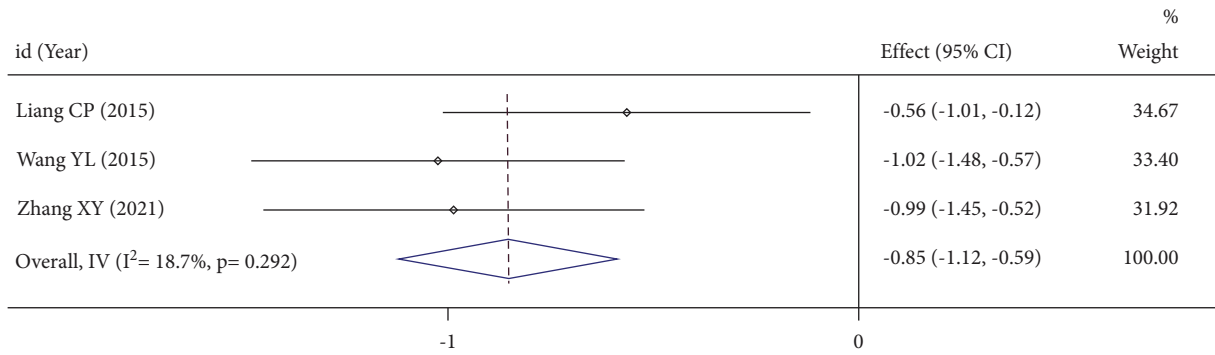
FIGURE 5: Forest plot of combined results for serum CA125 level.
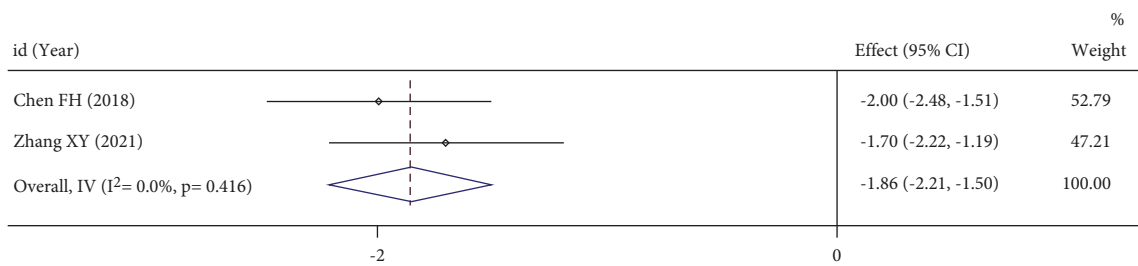
FIGURE 6: Forest plot of combined results for the VAS score of dysmenorrhea.

anti-inflammatory drugs to inhibit the synthesis of prostaglandins, inhibit lymphocyte activity and differentiation of activated T cells, reduce the stimulation of afferent nerve endings, act directly on nociceptors, and inhibit the formation and release of pain-inducing substances; oral contraceptives inhibit ovulation; highly effective progestins cause endometrial decidual-like changes, ultimately leading to endometrial atrophy, while negative feedback inhibits the hypothalamic-pituitary-ovarian axis; GnRH-a can downregulate pituitary function, cause temporary medical castration and estrogen status in the body, and can also bind to GnRH-a receptors in the periphery to inhibit the activity of eutopic and ectopic endometrial cells; gestrinone reduces ER and PR levels, reduces estrogen levels in the blood, and

reduces sex hormone binding globulin levels. Lesion resection, as one of the main methods to treat endometriosis, has the characteristics of good radical effect and low recurrence rate compared with drug therapy. Laparoscopic resection, as a minimally invasive surgery, compared with traditional laparotomy, laparoscopic surgery can significantly reduce the interference with the internal environment of patients, and at the same time, it can significantly reduce the fear and tension of patients to the operation, which is conducive to the postoperative recovery of patients. However, some studies have successively found that although laparoscopic resection has more advantages compared with traditional surgery, some patients still experience recurrence and even infertility [18].
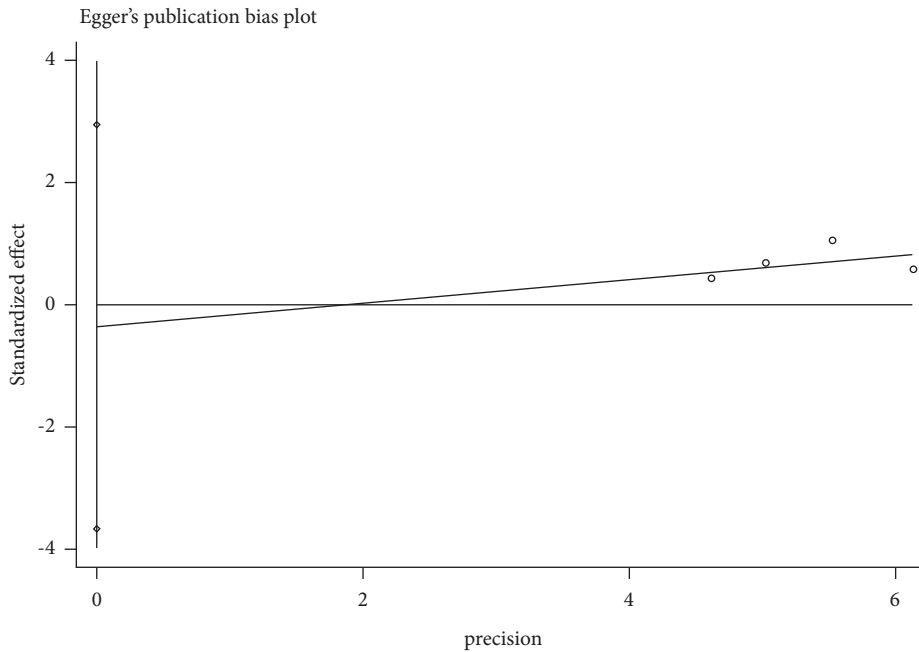
Figure 7: Egger's test.

At present, with the continuous progress of interventional therapy techniques, uterine artery embolization can also be used as one of the treatment options for endometriosis. In this study, a meta-analysis of 7 included RCT studies showed that there were significant differences in dysmenorrhea VAS score and serum CA125 level, both suggested that the results of the UAE group were better than those of the control group, and serum CA125, as a common serum tumor marker, belonged to the tumor-associated antigen recognized by the monoclonal antibody OC125 for epithelial ovarian cancer and was mostly used in the diagnosis of ovarian cancer and the detection after treatment. After treatment, the CA125 level in the observation group was significantly decreased, indicating that the diseased endometrium was better improved, and the VAS score of dysmenorrhea could directly reflect the improvement of the efficacy of patients after surgery, suggesting that the patients in the observation group could relieve more pain after surgery. However, the total effective rate and serum sex hormone levels were not significantly lower than those in the control group. Zhang Xinyu et al. [13] reported the occurrence of infection, bleeding, and other adverse reactions after UAE, but there was no statistical significance. As there were few reports related to overall complications, no meta-analysis was performed. It could not confirm whether the report of this complication was related to individual differences in the condition and physician's operation level.

The limitations of this paper include the following: ① 7 randomized controlled trials included have certain risks during the randomization process, which may cause certain selectivity bias; ② no relevant English literatures were retrieved, and only the Chinese literatures that met the criteria were studied; ③ the quality of the relevant literatures published at present is low, which may have a certain impact on the reliability of the results. The impact of commonly used antibiotics, e.g., amoxicillin [19], ornidazole [20], etc., on patient treatment also needs to be investigated.

To sum up, UAE is superior to traditional treatment in the postoperative serum CA125 level and VAS score of dysmenorrhea, but it still cannot replace medication and traditional surgery. In clinic, appropriate treatment should be selected according to the specific condition of patients with endometriosis. In the future, more large-scale randomized controlled trials with higher quality and higher sample size should be conducted to further study the treatment of endometriosis with UAE.

Data Availability

The data used in this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] C. Chapron, L. Marcellin, and B. Borghese, "Rethinking mechanisms, diagnosis and management of endometriosis," *Nature Reviews Endocrinology*, vol. 15, no. 11, pp. 666–682, 2019.

[2] J. H. Lang, "History, status and development of endometriosis," *Chinese Journal of Practical Gynecology and Obstetrics*, vol. 36, no. 3, pp. 193–196, 2020.

[3] S. As-Sanie, R. Black, and L. C. Giudice, "Assessing research gaps and unmet needs in endometriosis," *American Journal of Obstetrics and Gynecology*, vol. 221, no. 2, pp. 86–94, 2019.

[4] L. C. Giudice, S. As-Sanie, and J. C. A. Ferreira, "Once daily oral relugolix combination therapy versus placebo in patients with endometriosis-associated pain: two replicate phase 3, randomised," *double-blind, studies (SPIRIT 1 and 2), The Lancet*, vol. 399, no. 10343, pp. 2267–2279, 2022.

[5] Endometriosis Professional Committee of Obstetrics and Gynecologists Branch of Chinese Medical Doctor Association, "Chinese expert consensus on long-term management of endometriosis," *Chinese Journal of Obstetrics and Gynecology*, vol. 53, no. 12, pp. 836–841, 2018.

[6] T. Lu, C. Lu, H. Li et al., "The reporting quality and risk of bias of randomized controlled trials of acupuncture for migraine: methodological study based on STRICTA and RoB 2.0," *Complementary Therapies in Medicine*, vol. 52, Article ID 102433, 2020.

[7] C. P. Liang, "Observation on the efficacy of different surgical methods in the treatment of adenomyosis," *Henan Journal of Surgery*, vol. 21, no. 04, pp. 83-84, 2015.

[8] C. J. Lu and R. Wang, "Efficacy and safety analysis of laparoscopic resection combined with arterial embolism in the treatment of ovarian endometriosis," *Chinese Health Standard Management*, vol. 6, no. 30, pp. 53-54, 2015.

[9] K. Yuan, J. L. Zhang, and J. Y. Yan, "Uterine artery embolization with small-sized particles for the treatment of symptomatic adenomyosis: A 42-month clinical follow-up, International Journal of General," *Medicine*, vol. 14, no. 3575, 2021.

[10] F. H. Chen, Q. Wang, and Y. Deng, "Efficacy of uterine artery embolization in the treatment of severe dysmenorrhea caused by adenomyosis," *Journal of Clinical Psychosomatic Disorders*, vol. 24, no. 5, pp. 180–182, 2018.

[11] W. J. Walker and J. P. Pelage, "Uterine artery embolisation for symptomatic fibroids: clinical results in 400 women with imaging follow up," *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 109, no. 11, pp. 1262–1272, 2002.

[12] M. Popovic, S. Puchner, and D. Berzaczy, "Uterine artery embolization for the treatment of adenomyosis: a review," *Journal of Vascular and Interventional Radiology*, vol. 22, no. 7, pp. 901–909, 2011.

[13] X. Y. Zhang, "Comparison of the effects of uterine artery embolization and laparoscopic lesion resection in patients with adenomyosis," *China Minkang Medicine*, vol. 33, no. 3, pp. 129–131, 2021.

[14] W. T. Ross, J. M. Newell, and R. Zaino, "Appendiceal endometriosis: is diagnosis dependent on pathology evaluation? A prospective cohort study," *Journal of Minimally Invasive Gynecology*, vol. 27, no. 7, pp. 1531–1537, 2020.

[15] J. Zhou, L. He, and P. Liu, "Outcomes in adenomyosis treated with uterine artery embolization are associated with lesion vascularity: a long-term follow-up study of 252 cases," *PLoS One*, vol. 11, no. 11, 2016.

[16] L. K. Symons, J. E. Miller, V. R. Kay et al., "The immunopathophysiology of endometriosis," *Trends in Molecular Medicine*, vol. 24, no. 9, pp. 748–762, 2018.

[17] B. G. Patel, E. E. Lenk, D. I. Lebovic, Y. Shu, J. Yu, and R. N Taylor, "Pathogenesis of endometriosis: i," *Best Practice & Research Clinical Obstetrics & Gynaecology*, vol. 50, pp. 50–60, 2018.

[18] J. Wei and J. X. Li, "Comparison of the application value of open surgery and laparoscopic surgery in endometriosis," *Chinese Health Standards Administration*, vol. 5, no. 22, pp. 108-109, 2014.

[19] X. Cheng, F. Huang, K. Zhang, X. Yuan, and C Song, "Effects of none-steroidal anti-inflammatory and antibiotic drugs on the oral immune system and oral microbial composition in rats," *Biochemical and Biophysical Research Communications*, vol. 507, no. 1-4, pp. 420–425, 2018.

[20] X. Cheng, F. He, M. Si, P. Sun, and Q. Chen, "Effects of antibiotic use on saliva antibody content and oral microbiota in Sprague Dawley rats," *Frontiers in Cellular and Infection Microbiology*, vol. 12, p. 42, 2022.

Review Article

Kangaroo Care for Relieving Neonatal Pain Caused by Invasive Procedures: A Systematic Review and Meta-Analysis

Yunan Zhao, Yanjun Dong, and Jie Cao [ID]

Department of Neonatology, The First Hospital of China Medical University, Shenyang, Liaoning 110001, China

Correspondence should be addressed to Jie Cao; cjie880810@163.com

Objective. Neonates develop significant pain responses during invasive procedures, and nonpharmacological interventions are better means of pain relief. An increasing number of studies have confirmed the effectiveness of kangaroo care (KC) in relieving neonatal pain caused by invasive procedures, but conclusions are inconsistent. In this study, a literature search and meta-analysis were performed to evaluate the effect of kangaroo care on relieving neonatal pain. *Methods*. The works of literature related to the application of KC in neonatal invasive procedures in the databases of Pubmed, Embase, Springer Link, Ovid, CNKI, and CBM were searched, and the RCT literature from database establishment to July 2022, was selected to evaluate the risk of bias, combined with statistical pain relief outcome indicators. *Results*. 12 pieces of literature were finally included in this study, with a total of 1172 newborns, including 585 newborns (49.9%) using KC and 587 newborns (50.1%) using the control group method. Meta-analysis showed that an infant's heart rate during invasive procedures under KC intervention was significantly lower than that of other interventions ($MD = -6.77$, 95% CI (-13.03, -0.50), $Z = -2.12$, $P = 0.03$), but compared to other nonpharmacological interventions, there was no clear advantage in the overall evaluation of pain reduction in infants ($MD = -0.36$, 95% CI (-0.80, 0.08), $Z = -1.60$, $P = 0.11$). *Conclusion*. The heart rate of KC intervention during invasive procedures in infants is significantly lower than that of other interventions, and it can significantly relieve pain in infants, but the effect is not more than that of oral sucrose (or glucose) or standard care. KC combined with oral sucrose may achieve a better pain relief effect in infants, but more studies are still needed to verify it.

1. Introduction

Compared with older children and adults, newborns are more sensitive, intense, diffuse, and persistent in the perception of pain, and pain can cause significant physiological reactions, allergic reactions, and even chronic pain syndrome, and long-term physical discomfort, and lead to a series of short-term and long-term adverse effects such as difficulty concentrating, anxiety, cognitive behavioral disorders, poor adaptability, and growth retardation in childhood [1–3]. Neonatal pain is mainly derived from various invasive procedures during hospitalization (including heel blood sampling, arteriovenous puncture, intramuscular injection, and lumbar puncture) [4]. Different from adults, studies have reported that 80% to 90% of analgesic drugs have different degrees of adverse reactions in newborns [5], so it is difficult to widely use drugs to intervene in neonatal

pain in clinical practice. At present, there are more non-pharmacological intervention methods to relieve the symptoms of newborns, such as nonnutritive sucking, oral sucrose (or glucose), etc. [6]. Kangaroo care (KC), also known as skin-to-skin care, aims to provide humanized care for newborns, reduce the separation time between children and parents, promote closer mother-child relationships, make parents more confident in coping with low birth weight, and continuously improve the physical and behavioral stability of infants [7]. Increasing studies have confirmed the effectiveness of KC in relieving neonatal pain caused by invasive procedures, and compared with drug intervention, it has the characteristics of less risk, simplicity, and economic applicability. Although many literature studies on relieving neonatal pain caused by invasive procedures through KC have emerged at home and abroad in recent years, the conclusions obtained are not completely

consistent due to insufficient sample size, poor systematicness and integrity of design, and different results in some single-randomized controlled trials (RCTs). In a study by Campbell-Yeo et al. [8], 81 neonates randomized to KC were compared with 81 neonates treated with oral sucrose, and the results showed that there was no significant difference in the neonatal pain score (PIPP, premature infant pain profile score) after heel blood sampling compared with the control group treated with oral sucrose. However, in another randomized controlled study by de Sousa Freire et al. [9], newborns who also received KC intervention had lower PIPP scores after heel prick compared with newborns who received sucrose intervention, and the difference was statistically significant. The conclusions of such studies are completely different, and it is difficult to safely and effectively use kangaroo care to guide the clinical relief of neonatal pain. Therefore, this study aims to combine and analyze the relevant published literature studies by meta-analysis, in order to objectively evaluate the effect of KC on relieving neonatal pain and provide strong empirical evidence for the management of clinical neonatal pain.

2. Materials and Methods

2.1. Literature Database.
In this study, Pubmed, Embase, Springer Link, and Ovid were selected as English database sources, CNKI and CBM were selected as Chinese database sources, and the publication time of these pieces of literature was from the beginning of database establishment to July 2022. Two colleagues searched the pieces of literature published in the database with the keywords ("*Kangaroo nursing*" OR "*Kangaroo care*" OR "*skin to skin care*") AND ("*premature infant*") AND ("*pain*") AND ("*RCT*" OR "*Randomized controlled trial*"). Generally speaking, keywords with similar significance were connected with "*OR*," while those without similarity were connected with "*AND*."

2.2. Literature Inclusion and Exclusion Criteria

2.2.1. Inclusion Criteria.
The inclusion criteria were as follows: ① All pieces of literatures were single-center or multicenter-randomized controlled trials (*RCTs*); ② all subjects were infants born within one month, including full-term infants or premature infants, during hospitalization, the child underwent invasive procedures due to screening, diagnosis, and treatment of diseases (such as heel blood sampling, venipuncture, and intramuscular injection); ③ all studies were divided into the experimental group and the control group according to the randomization method for intervention, and we will perform a quality evaluation for the randomization process of each study. Children in the experimental group were intervened with kangaroo care (or skin-to-skin care) when undergoing invasive procedures, and infants were held in a comfortable position by their parents (specific measures varied depending on each study); in the control group, standard care or minimal care or other intervention methods (such as oral sucrose/glucose) were used [10].

2.2.2. Exclusion Criteria.
The exclusion criteria were as follows: ① we excluded non-RCT studies (such as controlled clinical studies, observational studies, and review studies); ② studies containing analgesics or sedatives within 24 to 72 hours before invasive procedures, or critically ill neonates unresponsive to painful stimuli; ③ studies which were unable to obtain outcome indicators or unable to obtain data on outcome indicators; ④ combined interventions, such as KC combined with sucrose; ⑤ studies that did not compare KC with other interventions for the purpose of the study, such as the comparison of the effect of KC performed by the father and mother; ⑥ literature studies in which the implementation scenario is specified as invasive operation in the study, such as the literature in the study in which the implementation scenario is lactation.

2.3. Literature Quality Evaluation.
Cochrane risk of bias V2.0 [11] was used to perform a quality evaluation for the included randomized controlled studies. This standard includes 5 domains (randomization process, implementation bias, data bias, data measurement bias, and selection bias) and 1 overall bias assessment. "Low risk," "some concerns of risk," and "high risk" were used to evaluate each domain (or overall bias).

2.4. Outcome Indicators.
Primary outcome indicators included the following indicators: (a) premature infant pain profile score (PIPP) [12]; (b) infant heart rate at manipulation; (c) infant oxygen saturation at manipulation. Although there are many scales to measure pain in newborns, we only take the most applied one, that is, the PIPP scale, which scores infant's responses to invasive procedures from fetal age, behavioral status, highest heart rate, oxygen desaturation, frowning, squeezing eyes, and deepening nasolabial folds, with higher scores indicating higher pain. The measurement time points are 2 to 5 min after invasive response, if the measurement data of multiple time points are reported in the literature, only the data of 2 min will prevail. If too few articles were included, we only performed descriptive analysis for indicators.

2.5. Literature Screening.
The pieces of literature retrieved from the database were screened by two colleagues, after the software deduplication, the abstract part was read, and the remaining pieces of literature were obtained after preliminary screening according to the previously established inclusion and exclusion criteria after full text reading, and finally, after quality evaluation, the pieces of literature with serious bias and low quality were removed, and the studies were finally included in the analysis.

2.6. Data Extraction.
Data extraction was performed by the researcher by reading the full text, and data such as interventions, the total number of people, grouping, characteristics of study subjects, and outcome indicators were extracted and entered into Excel sheets.

Identification of studies via databases and registers

Identification

Databases (n = 365):
    Pubmend (n = 106)
    EMBASE (n = 98)
    Spinger Link (n = 34)
    Ovid (n = 48)
    CNKI (n = 46)
    CBM (n = 33)

Duplicate records removed manually (n = 111)

Screening

Studies screened (n = 254)

Records excluded (total : 134)
    A) Not an RCT (n = 56)
    B) Not infant subject (n = 35)
    C) No kangaroo care intervention (n = 42)
    D) No control group (n = 33)

Studies sought for retrieval (n = 120)

Records not retrieved (n = 53)

Studies assessed for eligibility (n = 67)

Records excluded (n = 55)
    A) No available outcomes (n = 33)
    B) No data (n = 22)

Included

Studies included in meta-anlaysis (n = 12)

FIGURE 1: Literature selection flowchart.

2.7. Statistical Methods. (a) Continuous data (PIPP score, HR, and SPO2) used pooled MD with 95% CI as effect size and as the random-effect model, $P < 0.05$ was considered statistically significant, and the combined results were presented on forest plots; (b) for literature heterogeneity, Tau values were calculated, using Q check, and $P < 0.05$ was used to indicate heterogeneity of the results; (c) if there was heterogeneity between literature studies, subgroup analysis was adopted to assess the impact factor, and impact analysis was used to diagnose the stability of the results; (d) publication bias was quantified using Egger's test and presented using contour-enhanced funnel plots.

3. Results

3.1. Literature Screening Process and Results. Literature identification, screening, and exclusion followed PRISMA's recommendations and are shown in Figure 1. In this study, 365 pieces of literature were initially searched, and these 365 pieces of literature retrieved from the database were deduplicated and screened; only 12 pieces of literature [8, 9, 13–22] were included in the final study.

3.2. Basic Characteristics and Patient Characteristics of the Included Pieces of Literature. A total of 1172 infants were included in this study, of whom 585 neonates (49.9%) used KC,

587 neonates (50.1%) used other methods, and 2 literatures (16.7%) were injected with HBV vaccine. In the control group, oral sucrose (or glucose) was used in 6 articles and standard care or minimal care in 6 articles. Basic characteristics of the literature are shown in Table 1.

3.3. Literature Bias and Quality Assessment. 5 (41.7%) of the 12 articles included in this study [13, 16, 18, 19, 22] had "some risk of bias" in terms of the randomization process, data measurement, intervention deviations, and selective bias, and the remaining 7 (58.3%) articles were "low risk" with a good overall quality. Summary of bias assessed according to Cochrane ROB 2.0 and details of the assessment are summarized in Figures 2(a) and 2(b).

3.4. Metaquantitative Analysis Results of Outcome Indicators

3.4.1. PIPP Score after the Procedure. Among the 12 included pieces of literature, 7 pieces of literature [8, 9, 13, 14, 17–19] tried to draw the comparison between KC and other nondrug intervention methods in invasive procedures for infants and reported the PIPP score indicators, 5 pieces of literature [8, 9, 14, 17, 18] were treated with sucrose intervention in the control group, and 2 pieces of literature [13, 19] were treated with standard care in the

TABLE 1: Basic characteristics, patient characteristics, and outcome indicators of the included pieces of literature.

Author	Year	Intention-to-treat total	Number of the experimental group	Number of the control group	GA (weeks)	Painful stimulus occasions	Experimental group method	Control group method	Outcome indicators
Campbell-Yeo et al. [8]	2019	162	81	81	32.6 ± 2.1	Heel lance	KC through the procedure	Oral Sucrose (1 ml, 24%)	(a)
de Sousa Freire et al. [9]	2008	62	31	31	33.7 ± 1.57	Heel lance	KC through the procedure	Oral glucose (1 ml, 25%)	(a)
Mitchell et al. [13]	2013	26	13	11	29	Nasal suctioning	2 hours daily skin-to-skin holding	Minimal parent holding	(a)
Sen et al. [14]	2020	64	32	32	34.38 ± 1.59	Heel lance	KC through the procedure	Oral Sucrose (0.5 ml, 24%)	(a) (b) (c)
Gao et al. [15]	2014	80	40	40	37	Heel lance	KC through the procedure	Standard care	(b)
Cong et al. [16]	2012	48	25	23	30 ± 2.2	Heel lance	KC through the procedure	Standard care	(b)
Nimbalkar et al. [17]	2020	200	100	100	33.6 ± 1.89	Heel lance	KC through the procedure	Oral Sucrose (1 ml, 24%)	(a) (c)
Shukla et al. [18]	2018	100	50	50	32.79 ± 2.34	Heel lance	KC through the procedure	Oral Sucrose (1 ml, 24%)	(a) (c)
Kristoffersen et al. [19]	2019	46	21	25	33 ± 2	Eye examinations	KC through the procedure	Standard care	(a)
Chermont et al. [20]	2009	320	160	160	39 ± 1	Intramuscular injection of hepatitis B vaccine	KC through the procedure	Oral glucose (1 ml, 25%)	(d)
Kostandy et al. [21]	2013	36	17	19	39.7 ± 1.29	Intramuscular injection of hepatitis B vaccine	KC through the procedure	Standard care	(b)
Gray et al [22]	2000	30	15	15	NA	Heel lance	KC through the procedure	Standard care	(b)

Note. PIPP, premature infant pain profile; GA, gestational age; KC, kangaroo care; HR, heart rate; SPO$_2$, pulse oxygen saturation; NFCS, neonatal facial coding system; and NA, not available. Outcomes: (a) PIPP score after the procedure; (b) HR value during the procedure; (c) SPO$_2$ during the procedure; (d) the NFCS Score.

control group. Combined results showed that 7 articles presented statistical heterogeneity ($Chi^2 = 24.58$, $df = 6$, $P < 0.01$), and the random-effect model was used to obtain a combined effect size ($MD = -0.36$, 95% CI (-0.80, 0.08), $Z = -1.60$, $P = 0.11$), indicating KC compared with other nonpharmacological interventions; there was no clear advantage for pain reduction in infants as shown in Figure 3.

3.4.2. Heart Rate during the Procedure. 5 articles [14–16, 21, 22] reported the comparison of the heart rate between KC and other nonpharmacological interventions in invasive procedures, and the combined results showed that 5 articles were statistically heterogeneous ($Chi^2 = 17.86$, $df = 4$, $P < 0.01$); the fixed-effect model was used to obtain a combined effect size ($MD = -6.77$, 95% CI (-13.03, -0.50), $Z = -2.12$, $P = 0.03$), indicating that infants under KC

(a)

(b)

FIGURE 2: Risk of bias for inclusion of 12 articles based on Cochrane risk of bias 2.0: (a) summary of the bias and (b) detail of the bias.

FIGURE 3: The PIPP score after the procedure: comparing KC with other interventions.

FIGURE 4: Heart rate during the procedure: comparing KC with other interventions.

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight (%)	Mean Difference IV, Random, 95% Cl
occasions = heel lance								
Campbell-Yeo M etal (8) 2019	5.25	2.2400	81	4.96	2.5100	81	14.9	0.29 [-0.44; 1.02]
de Sousa Freire NB et al (9) 2008	6.77	2.1000	31	7.94	2.5000	31	9.3	-1.17 [-2.32; -0.02]
Sen E et al (14) 2020	3.00	0.4400	32	4.00	0.6800	32	22.9	-1.00 [-1.28; -0.72]
Nimbalkar S et al (17) 2020	5.98	3.3400	100	6.84	3.5900	100	11.5	0.14 [-0.82; 1.10]
Shukla V et al (18) 2018	7.74	2.4300	50	8.10	2.8200	50	10.6	-0.36 [-1.39; 0.67]
Total (95% Cl)			294			294	69.3	-0.44 [-1.04; 0.16]

Heterogeneity: Tau^2 = 0.2994; Chi^2 = 15.13, df = 4 (P < 0.01); I^2 = 74%

occasions = nasal suctioning								
Mitchell AJ et al (13) 2013	7.64	0.4000	13	7.89	0.2100	11	23.4	-0.25 [-0.50; 0.00]

occasions = eye examinations								
Kristoffersen L et al (19) 2019	7.00	2.4000	21	6.80	2.3000	25	7.4	0.20 [-1.17; 1.57]
Total (95% Cl)			328			330	100.0	-0.36 [-0.80; 0.08]

Heterogeneity: Tau^2 = 0.2011; Chi^2 = 24.58, df = 6 (P < 0.01); I^2 = 76%
Test for subgroup differences: Chi^2 = 0.79, df = 2 (P = 0.68)

(a)

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight (%)	Mean Difference IV, Random, 95% Cl
Control = Oral Sugars								
Campbell-Yeo M etal (8) 2019	5.25	2.2400	81	4.96	2.5100	81	14.9	0.29 [-0.44; 1.02]
de Sousa Freire NB et al (9) 2008	6.77	2.1000	31	7.94	2.5000	31	9.3	-1.17 [-2.32; -0.02]
Sen E et al (14) 2020	3.00	0.4400	32	4.00	0.6800	32	22.9	-1.00 [-1.28; -0.72]
Nimbalkar S et al (17) 2020	5.98	3.3400	100	6.84	3.5900	100	11.5	0.14 [-0.82; 1.10]
Shukla V et al (18) 2018	7.74	2.4300	50	8.10	2.8200	50	10.6	-0.36 [-1.39; 0.67]
Total (95% Cl)			294			294	69.3	-0.44 [-1.04; 0.16]

Heterogeneity: Tau^2 = 0.2994; Chi^2 = 15.13, df = 4 (P < 0.01); I^2 = 74%

Control = Standard care								
Mitchell AJ et al (13) 2013	7.64	0.4000	13	7.89	0.2100	11	23.4	-0.25 [-0.50; 0.00]
Kristoffersen L et al (19) 2019	7.00	2.4000	21	6.80	2.3000	25	7.4	0.20 [-1.17; 1.57]
Total (95% Cl)			34			36	30.7	-0.24 [-0.48; 0.01]

Heterogeneity: Tau2 = 0; Chi2 = 0.4, df = 1 (P < 0.53); I2 = 0%

| Total (95% Cl) | | | 328 | | | 330 | 100.0 | -0.36 [-0.80; 0.08] |

Heterogeneity: Tau^2 = 0.2011; Chi^2 = 24.58, df = 6 (P < 0.01); I^2 = 76%
Test for subgroup differences: Chi^2 = 0.38, df = 1 (P = 0.54)

(b)

FIGURE 5: Subgroup analysis according to (a) occasions or (b) control methods.

intervention had significantly smaller heart rates during invasive procedures than those in other interventions as shown in Figure 4.

3.4.3. Heterogeneity Investigation.
In the combined analysis of the PIPP score, heterogeneity between articles was statistically significant, and all 7 articles were divided into subgroups according to the scenario of intervention application and according to the way of intervention in the control group. However, heterogeneity between subgroups was not statistically significant ($P = 0.68$; $P = 0.54$), which illustrated the scenario of intervention application, and the mode of intervention in the control arm was not a source of heterogeneity as shown in Figures 5(a) and 5(b).

3.4.4. Influence Analysis.
In the influence analysis on PIPP score outcome indicators, it was found that all pieces of literature were within the acceptable range, indicating a good stability, as shown in Figure 6.

3.4.5. Publication Bias Analysis.
In the combined analysis of PIPP score outcome indicators, publication bias was measured using Egger's test for the results, $t = 0.66$, $P = 0.54$, indicating that there was no publication bias, and the contour-enhanced funnel plot is shown in Figure 7.

4. Discussion

Although newborns are unable to express pain in words, they have perceptual instinct to negative stimuli such as pain and can show significant physiological responses, such as increased heart rate, increased blood pressure, decreased oxygen saturation, and increased respiratory rate, and neonatal pain can be comprehensively evaluated with the help of behavioral responses such as crying and duration of painful facies[23]. Therefore, the PIPP score, heart rate, and oxygen saturation indexes were selected in this study to comprehensively evaluate the effect of kangaroo care on relieving neonatal pain caused by invasive procedures.
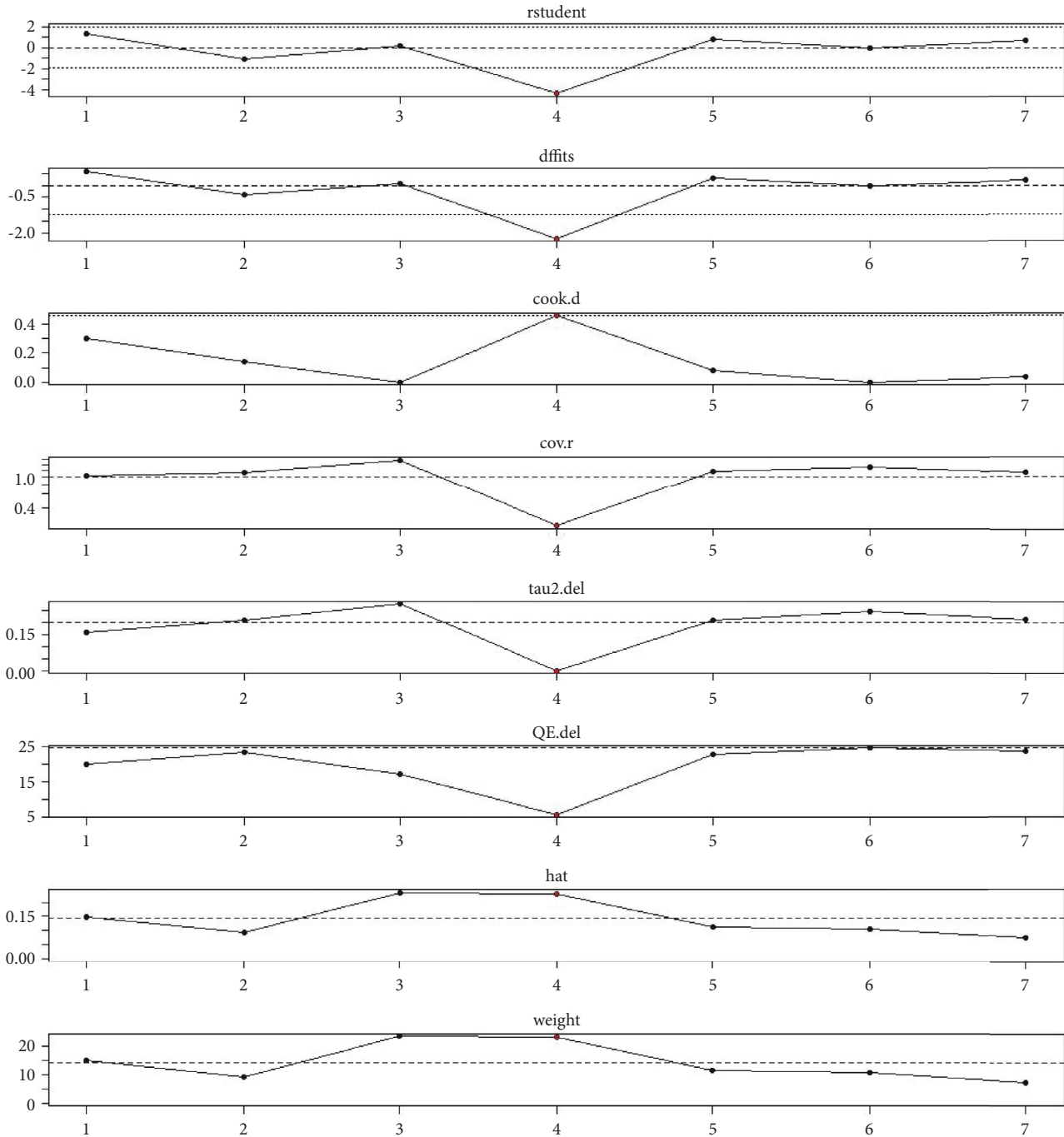
Figure 6: Influence analysis.

Among the 12 pieces of literature included in this study, only 7 pieces of literature reported the PIPP score, 6 pieces of literature reported the heart rate, and 1 literature reported oxygen saturation. Meta-analysis showed that KC had no significant advantage over other nondrug interventions in the overall assessment of infant pain; however, infants' heart rate during invasive procedures with KC intervention was significantly lower than in other interventions. Because there was only one literature reporting oxygen saturation, a meta-analysis was not performed at that time. Although KC is not more effective in reducing pain than oral sucrose (or glucose) or standard care, it remains positive for neonatal pain relief, which may be due to kangaroo care inhibiting the activity of the hypothalamic-pituitary-adrenocortical axis, reducing salivary cortisol, serum cortisol, and β-endorphin secretion levels, and stimulating C-afferent fibers to excite the limbic system of the brain to produce a sense of pleasure and inhibit the conduction of pain signals, thereby reducing operant pain and resulting in a decreased heart rate [24–26]. However, also as a nonpharmacological intervention, oral sucrose (or glucose) is also an important intervention, and sucrose and glucose are the most commonly used
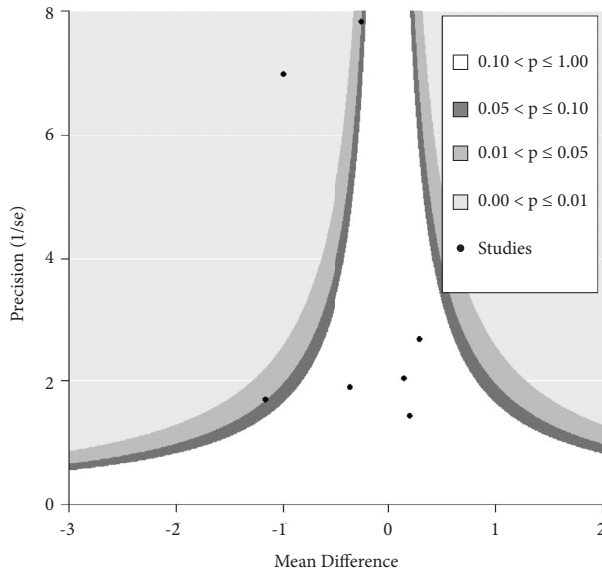
Figure 7: Contour-enhanced funnel plot.

sweeteners, which are effective and simple to use, and have no documented side effects and have gained wide use in the clinical setting for management of neonatal pain [27, 28]. In the literature [8], although there was no significant difference in PIPP scores between KC alone and oral sucrose, the investigators combined oral sucrose with KC in 81 neonates, and compared with oral sucrose, PIPP scores in KC were significantly lower than those in the oral sucrose group, which suggests that KC combined with oral sucrose is coapplied in children to reduce pain during invasive procedures, and this may lead to better results. This was mapped in literature [20]; however, as this was not the purpose of this study, it could be left for verification by subsequent studies.

In literature [16], the investigators randomized two experimental groups, in which KC started 30 min or 15 min before neonatal heel blood sampling, and the results showed no difference in pain relief between the two groups, which showed that KC had little effect when it started, and it was important that maternal contact and comfort were provided for neonatal pain stimulation to reduce pain and anxiety. In a study by Shukla et al. [29], comparing the effect of KC performed by fathers and mothers separately, the results showed no significant difference, which may be related to the fact that newborns are not sensitive to the giver of KC.

In this meta-analysis, heterogeneity in pieces of literature was obvious, but we performed subgroup analysis according to factors that may cause heterogeneity of the literature (scenario of intervention application and the mode of intervention in the control group), and there was no significant difference between the groups. This suggested that the source of heterogeneity is not related to the scenario of intervention application, and the mode of intervention in the control group may be related to different characteristics of newborns, different strategies implemented by KC, and confounding factors such as the number of included sample sizes.

Although the quality evaluation by Cochrane ROB 2.0 of 12 articles suggested that 7 articles had "some concerns of risk," the overall quality of the articles was good, the results were stable, and there was no publication bias. Because different pieces of literature had diversified reporting indicators for pain, for example, some pieces of literature [20] adopted the NFCS score to report pain, and many other studies [30, 31] were excluded because there were no available data, which makes this analysis not comprehensive. For this topic, more studies with better quality still need to be included from different perspectives with different indicators for in-depth analysis. On the other hand, the impact of conventional NSAIDs or antibiotics, e.g., aspirin [32], amoxicillin [33], etc., on neonatal health also deserves attention.

The heart rate of KC intervention during invasive procedures in infants was significantly lower than that of other interventions, significantly relieving pain in infants, but the effect did not exceed that of oral sucrose (or glucose) or standard care, and the combination of KC and oral sucrose may have a better effect on pain relief in infants, but more studies are still needed to verify this effect.

Data Availability

The data used in 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] D. Koller and R. D. Goldman, "Distraction techniques for children undergoing procedures: a critical review of pediatric research," *Journal of Pediatric Nursing*, vol. 27, no. 6, pp. 652–681, 2012.

[2] G. Pancekauskaitė and L. Jankauskaitė, "Paediatric pain medicine: pain differences, recognition and coping acute procedural pain in paediatric emergency room," *Medicina (Kaunas)*, vol. 54, no. 6, p. 94, 2018.

[3] D. Harrison, J. Reszel, M. Bueno et al., "Breastfeeding for procedural pain in infants beyond the neonatal period," *Cochrane Database of Systematic Reviews*, vol. 10, no. 10, Article ID CD011248, 2016.

[4] B. Stevens, C. Johnston, P. Petryshen, and A. Taddio, "Premature infant pain profile: development and initial validation," *The Clinical Journal of Pain*, vol. 12, no. 1, pp. 13–22, 1996.

[5] M. Filippa, M. G. Monaci, C. Spagnuolo, P. Serravalle, R. Daniele, and D. Grandjean, "Maternal speech decreases pain scores and increases oxytocin levels in preterm infants during painful procedures," *Scientific Reports*, vol. 11, no. 1, Article ID 17301, 2021.

[6] C. McNair, M. Campbell-Yeo, C. Johnston, and A. Taddio, "Nonpharmacologic management of pain during common needle puncture procedures in infants: current research evidence and practical considerations: an update," *Clinics in Perinatology*, vol. 46, no. 4, pp. 709–730, 2019.

[7] H. Zengin and N. Cinar, "Designing dress (Sarbebe) for kangaroo care, the effect of kangaroo care provided with this

dress on mother and newborn's comfort," *Health Care for Women International*, vol. 43, no. 6, pp. 642–662, 2022.

[8] M. Campbell-Yeo, C. C. Johnston, B. Benoit et al., "Sustained efficacy of kangaroo care for repeated painful procedures over neonatal intensive care unit hospitalization: a single-blind randomized controlled trial," *Pain*, vol. 160, no. 11, pp. 2580–2588, 2019.

[9] N. B. de Sousa Freire, J. B. S. Garcia, and Z. C. Lamy, "Evaluation of analgesic effect of skin-to-skin contact compared to oral glucose in preterm neonates," *Pain*, vol. 139, no. 1, pp. 28–33, 2008.

[10] B. Stevens, J. Yamada, and A. Ohlsson, "Sucrose for analgesia in newborn infants undergoing painful procedures," *Cochrane Database of Systematic Reviews*, vol. 3, Article ID CD001069, 2004.

[11] S. Minozzi, K. Dwan, F. Borrelli, and G. Filippini, "Reliability of the revised Cochrane risk-of-bias tool for randomised trials (RoB2) improved with the use of implementation instruction," *Journal of Clinical Epidemiology*, vol. 141, pp. 99–105, 2022.

[12] B. J. Stevens, S. Gibbins, J. Yamada et al., "The premature infant pain profile-revised (PIPP-R): initial validation and feasibility," *The Clinical Journal of Pain*, vol. 30, no. 3, pp. 238–243, 2014.

[13] A. J. Mitchell, C. C. Yates, D. K. Williams, J. Chang, and R. W. Hall, "Does daily kangaroo care provide sustained pain and stress relief in preterm infants?" *Journal of Neonatal-Perinatal Medicine*, vol. 6, no. 1, pp. 45–52, 2013.

[14] E. Sen and G. Manav, "Effect of kangaroo care and oral sucrose on pain in premature infants: a randomized controlled trial," *Pain Management Nursing*, vol. 21, no. 6, pp. 556–564, 2020.

[15] H. Gao, G. Xu, H. Gao et al., "Effect of repeated Kangaroo Mother Care on repeated procedural pain in preterm infants: a randomized controlled trial," *International Journal of Nursing Studies*, vol. 52, no. 7, pp. 1157–1165, 2015.

[16] X. Cong, R. M. Cusson, S. Walsh, N. Hussain, S. M. Ludington-Hoe, and D. Zhang, "Effects of skin-to-skin contact on autonomic pain responses in preterm infants," *The Journal of Pain*, vol. 13, no. 7, pp. 636–645, 2012.

[17] S. Nimbalkar, V. V. Shukla, V. Chauhan et al., "Blinded randomized crossover trial: skin-to-skin care vs. sucrose for preterm neonatal pain," *Journal of Perinatology*, vol. 40, no. 6, pp. 896–901, 2020.

[18] V. Shukla, A. Chapla, J. Uperiya, A. Nimbalkar, A. Phatak, and S. Nimbalkar, "Sucrose vs. skin to skin care for preterm neonatal pain control-a randomized control trial," *Journal of Perinatology*, vol. 38, no. 10, pp. 1365–1369, 2018.

[19] L. Kristoffersen, R. Støen, H. Bergseng et al., "Skin-to-skin contact during eye examination did not reduce pain compared to standard care with parental support in preterm infants," *Acta Paediatrica*, vol. 108, no. 8, pp. 1434–1440, 2019.

[20] A. G. Chermont, L. F. M. Falcão, E. H. L. de Souza Silva, R. de Cássia Xavier Balda, and R. Guinsburg, "Skin-to-skin contact and/or oral 25% dextrose for procedural pain relief for term newborn infants," *Pediatrics*, vol. 124, no. 6, pp. e1101–e1107, 2009.

[21] R. Kostandy, G. C. Anderson, and M. Good, "Skin-to-skin contact diminishes pain from hepatitis B vaccine injection in healthy full-term neonates," *Neonatal Network*, vol. 32, no. 4, pp. 274–280, 2013.

[22] L. Gray, L. Watt, and E. M. Blass, "Skin-to-skin contact is analgesic in healthy newborns," *Pediatrics*, vol. 105, no. 1, p. e14, 2000.

[23] X. Cong, S. M. Ludington-Hoe, G. McCain, and P. Fu, "Kangaroo Care modifies preterm infant heart rate variability in response to heel stick pain: pilot study," *Early Human Development*, vol. 85, no. 9, pp. 561–567, 2009.

[24] S. M. Ludington-Hoe, R. Hosseini, and D. L. Torowicz, "Skin-to-skin contact (Kangaroo Care) analgesia for preterm infant heel stick," *AACN Clinical Issues: Advanced Practice in Acute and Critical Care*, vol. 16, no. 3, pp. 373–387, 2005.

[25] T. C. Castral, F. Warnock, A. M. Leite, V. J. Haas, and C. G. Scochi, "The effects of skin-to-skin contact during acute pain in preterm newborns," *European Journal of Pain*, vol. 12, no. 4, pp. 464–471, 2008.

[26] F. Okan, A. Ozdil, A. Bulbul, Z. Yapici, and A. Nuhoglu, "Analgesic effects of skin-to-skin contact and breastfeeding in procedural pain in healthy term neonates," *Annals of Tropical Paediatrics*, vol. 30, no. 2, pp. 119–128, 2010.

[27] J. J. Liaw, L. Yang, K. W. Katherine Wang, C. M. Chen, Y. C. Chang, and T. Yin, "Non-nutritive sucking and facilitated tucking relieve preterm infant pain during heel-stick procedures: a prospective, randomised controlled crossover trial," *International Journal of Nursing Studies*, vol. 49, no. 3, pp. 300–309, 2012.

[28] R. R. Pillai Riddell, N. M. Racine, H. G. Gennis et al., "Non-pharmacological management of infant and young child procedural pain," *Cochrane Database of Systematic Reviews*, vol. 2015, no. 12, Article ID CD006275, 2015.

[29] V. V. Shukla, A. J. Chaudhari, S. M. Nimbalkar, A. G. Phatak, D. V. Patel, and A. S. Nimbalkar, "Skin-to-Skin care by mother vs. Father for preterm neonatal pain: a randomized control trial (environ trial)," *International Journal of Pediatrics*, vol. 2021, pp. 1–6, 2021.

[30] D. V. Patel, S. N. Soni, V. V. Shukla et al., "Efficacy of skin-to-skin care versus swaddling for pain control associated with vitamin K administration in full-term neonates: a randomized controlled trial," *Journal of Tropical Pediatrics*, vol. 68, no. 4, 2022.

[31] Z. Kashaninia, F. Sajedi, M. Rahgozar, and F. A. Noghabi, "The effect of Kangaroo Care on behavioral responses to pain of an intramuscular injection in neonates," *Journal for Specialists in Pediatric Nursing*, vol. 13, no. 4, pp. 275–280, 2008.

[32] X. Cheng, F. Huang, K. Zhang, X. Yuan, and C. Song, "Effects of none-steroidal anti-inflammatory and antibiotic drugs on the oral immune system and oral microbial composition in rats," *Biochemical and Biophysical Research Communications*, vol. 507, no. 1-4, pp. 420–425, 2018.

[33] X. Cheng, F. He, M. Si et al., "Effects of antibiotic use on saliva antibody content and oral microbiota in Sprague Dawley rats [J]," *Frontiers in Cellular and Infection Microbiology*, p. 42, 2022.

Retraction

Retracted: Analysis of Signs and Effects of Surgical Breast Cancer Patients Based on Big Data 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] Z. Hong, Q. Xu, X. Yan, R. Zhang, Y. Ren, and Q. Tong, "Analysis of Signs and Effects of Surgical Breast Cancer Patients Based on Big Data Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3373553, 8 pages, 2022.

Research Article

Analysis of Signs and Effects of Surgical Breast Cancer Patients Based on Big Data Technology

Zhen Hong ⬤,[1] Qin Xu,[2] Xin Yan,[1] Ran Zhang,[1] Yuanfang Ren,[3] and Qian Tong[3]

[1]*School of Nursing, Jiangsu Jiankang Vocational College, Nanjing 211800, China*
[2]*School of Nursing, Nanjing Medical University, Nanjing 211166, China*
[3]*Department of Thyroid and Breast Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China*

Correspondence should be addressed to Zhen Hong; 19402379@masu.edu.cn

Big data in health care has gained popularity in recent years for disease prediction. Breast cancer infections are the most common cancer in urban Indian women, as well as women internationally, and are impacted by many events across countries and regions. Breast malignant growth is a notable disease among Indian women. According to the WHO, it represents 14% of all malignant growth tumors in women. A couple of studies have been directed utilizing big data to foresee breast malignant growth. Big data is causing a transformation in healthcare, with better and more ideal results. Monstrous volumes of patient-level data are created by using EHR (Electronic Health Record) systems data because of fast mechanical upgrades. Big data applications in the healthcare business will assist with improving results. Conventional forecast models, then again, are less productive in terms of accuracy and error rate because the exact pace of a specific calculation relies upon different factors such as execution structure, datasets (little or enormous), and kinds of datasets utilized (trait-based or picture based). This audit article looks at complex information mining, AI, and profound learning models utilized for recognizing breast malignant growth. Since "early identification is the way to avoidance in any malignant growth," the motivation behind this audit article is to support the choice of fitting breast disease expectation calculations, explicitly in the big information climate, to convey powerful and productive results. This survey article analyzes the precision paces of perplexing information mining, AI, and profound learning models utilized for distinguishing breast disease on the grounds that the exactness pace of a specific calculation relies upon different factors such as execution structure, datasets (little or enormous), and dataset types (quality based or picture based). The reason for this audit article is to aid the determination of suitable breast disease expectation calculations, explicitly in the big information climate, to convey successful and productive outcomes. Thus, "Early discovery is the way to counteraction in the event of any malignant growth."

1. Introduction

Big data has its roots from 1941 when the Oxford dictionary of English published the first mention to the concept of "information explosion." In a study published by the National Academy of Sciences in 1996, James Maar emphasized the concept of a "huge data set" (1). However, it was not until 1997 that the term "Big Data" was coined in an article in the Association for Computing Machinery's Digital Library (2), alluding to the technological issue of interpreting massive volumes of data. Since then, it has been used to refer to "structured or unstructured data whose extremely huge volume necessitates specialised analysis techniques." Over

the last decade, internet behemoths (Google, Amazon, Facebook, Apple, and Twitter) have built such technologies, ensuring a consistent marginal cost of data exploitation regardless of scale.

Today, Big Data is defined by the 5 vs of the data exploited: volume, velocity, variety, veracity, and value. The vast volumes and fast speed of data processing are the results of lower storage prices and increased computing capability [1]. The rising digitization of information media has resulted in a greater diversity of data (pictures, texts, databases, connected devices, and so on). Finally, the accuracy of the data, which determines the job's worth, is an important consideration for any automated data analysis

project [2–5].Indeed, if the data is huge, precise, and well-suited to the problem at hand, an algorithm can be tremendously strong. Multiplying sources and crossings without concern for data quality may result in incorrect results, particularly in the health area. The rise of "Open Data," which refers to data gathered and kept by multiple organizations and made available to individuals and businesses, has coincided with the rise of the internet development of big data [6].

The 5 vs, on the other hand, are insufficient to capture the core of big data's innovation. The knowledge of these algorithms is at the core of data scientists' work.

Breast cancer diagnosis and treatment have come a long way in the previous three decades. Individual or organized breast screening activities, as well as advances in breast imaging equipment, have all contributed to this breakthrough. Without a doubt, a sub-space of man-made consciousness known as "AI" permits designers to make algorithms equipped for collecting information and knowledge from tests without being directed by people or expressly modified to deal with a particular errand, bringing about their focal job in the data esteem chain, with the rest of to propels in careful methods or clinical medicines [7–11].

The advent of big data technology has recently piqued the curiosity of medical professionals concerned with breast health. Indeed, storage capacity has expanded significantly over the previous three decades, bringing about larger amounts and a more noteworthy variety of clinical information that has been saved (mammography filters, 3D ultrasound, MRI, genomic information, neurotic information, and so forth). As of not long ago, this information was every now and again utilized at a singular level over the course of time to produce a conclusion and a helpful arrangement, follow ailment improvement, and conjecture a patient's forecast. Besides, just coordinated information on a factual scale was utilized, which addressed just a tiny level of the open and useable data sources.

The rest of putting away information cemeteries was scarcely apparent to clinical professionals. The primary guarantee of great data is that it will empower the utilization of all information sources, even unstructured ones such as text-based patient reports or photographs, influencing clinical exploration and, at last, patient consideration [12, 13].

To completely appreciate how big data might improve breast health care, two developments must be examined. To start, the production environment that has arisen throughout the past ten years takes into account the execution of predefined procedures on immense volumes and various kinds of data. Second, AI algorithms and their programming language executions might gain from data to reveal examples and correlations, coming about in significant insights. Data researchers are seasoned veterans at dealing with these unique techniques [14].

The success of these cutting-edge investigations is dependent on transdisciplinarity. It takes time to develop a shared semantics between medical personnel and data scientists, and the breast disease units' experience working in transversal organizations will be invaluable in developing a framework for these interactions. Patients and, more broadly, civic society should be included in Big Data projects, in addition to medical professionals and data scientists, because only rigorous adherence to privacy standards can ensure their success and profitability [15, 16].

1.1. Breast Cancer.

Breast disease is a fiery growth that begins in the breast and spread to the remainder of the body. Disease emerges when cells start to multiply wildly. Breast malignant growth is a gathering of infections wherein breast tissue cells move and create uncontrolled, finishing in a protuberance or mass. Most breast tumors start in the dairy organs. Mammograms, breast self-assessment (BSE), biopsy, and refined breast tissue measurements are used to assess breast malignant growth. A medical operation, radiation, chemical treatment, chemotherapy, and laser treatment can all be used to treat breast cancer.

At the point when harmful development cells enter the circulatory or lymphatic frameworks and travel to various pieces of the body, breast ailment could spread. Breast cancerous growth cells typically structure an irregularity or cancer, which should be visible on an X-beam or felt as a firm mass. Monitoring controllable gamble variables can assist with lessening the probability of creating breast malignant growth [17]. Breast malignant growth is principally a female infection, but it is turning out to be more normal in men.

1.1.1. Types of Breast Cancer.

Obtrusive breast cancer and noninvasive breast cancer are two kinds of breast diseases that can be characterized based on regardless of whether the malignant growth has spread (Figure 1) [18].

Obtrusive breast disease, then again any sort of breast-threatening development that has spread to the breast tissues, is suggested as nosy ductal carcinoma (IDC) which attacks the incorporating tissues in the breasts, making them substantial and perceptible on mammograms. Intrusive breast disease represents 81% of all cases of invasive breast disease. Then again, any sort of dangerous breast development that has spread to the incorporating breast tissues is implied as meddlesome ductal carcinoma (IDC) attacks the including tissues of the breasts, making them discernible and perceptible on mammograms. Obtrusive breast malignant growth represents 81% of all cases. Invasive lobular carcinoma (ILC) infiltrates the breast tissues on both sides and is microscopically identified as inflammatory carcinoma, often known as IBC (fiery breast disease); it is an intriguing and perilous sort of breast malignant growth that appears as a rash or bothered skin region [19, 20].

It blocks the lymph vessels in the skin of the breasts. Since searing breast disease cannot be perceived by mammography or ultrasound, it ought to be found imperceptibly [21–24]. Paget's infection is an especially interesting sort of breast disease. Paget's disease of the breast starts on the areola and spreads to the encompassing dark circle of skin (areola). Intrusive ductal carcinoma is the medical term for male breast cancer in cutting-edge stages, Figure 2.
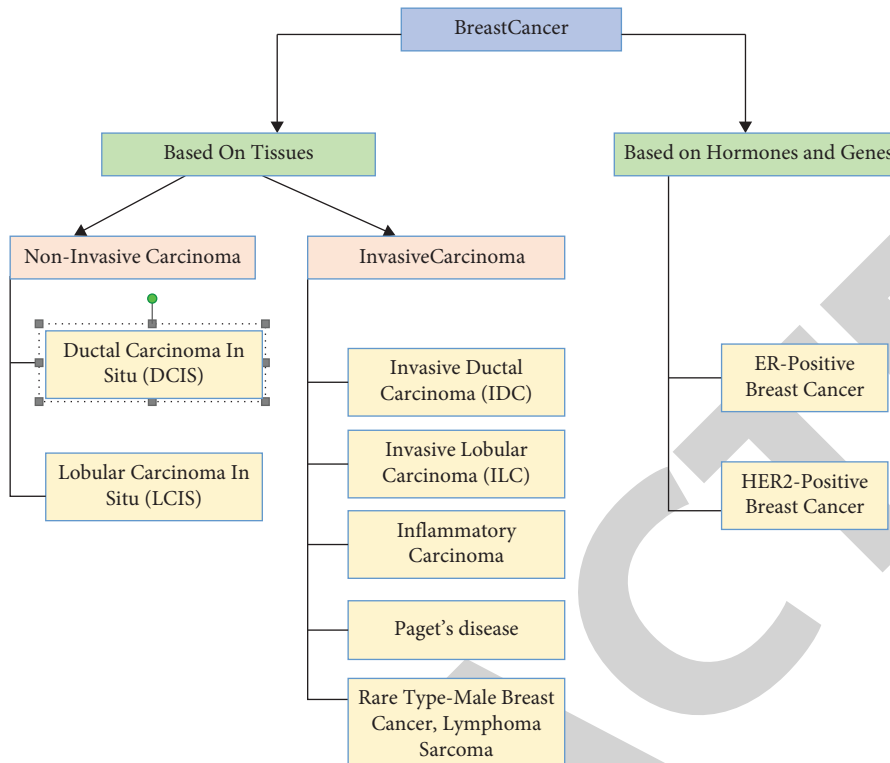
Figure 1: Classification of breast cancer.

Age standardized rate per 1,00,000 (*world*)
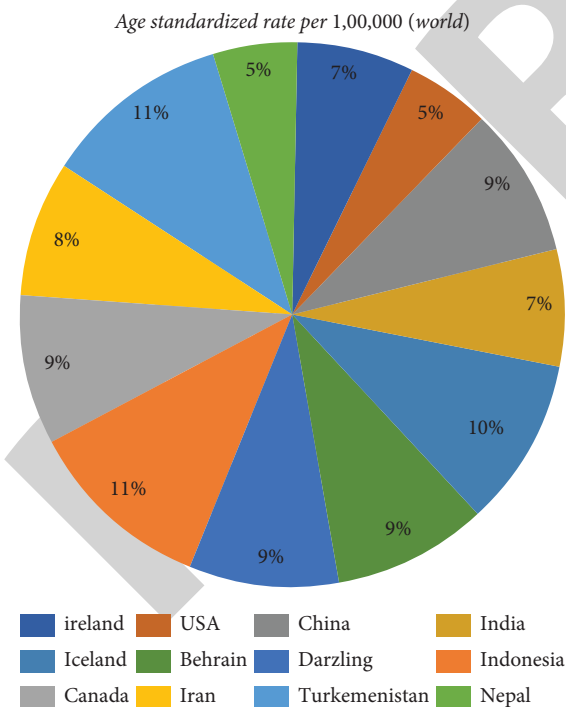
Figure 2: Preventing breast Cancer stats according WCRFI [20].

1.1.2. Breast Cancer Symptoms and Signs. Another lymph center point or hard mass around the Breast or underarm district is a truly common breast-threatening development sign. Breast dangerous development is consistently portrayed by a simple, hard knot with sporadic edges; however, it can likewise be a delicate, delicate mass. Breast malignant growth frequently has no side effects in light of the fact that the cancer is little and simple to treat. Consequently, screening is vital for right on-time identification. The signs and symptoms of breast cancer vary from person to person. Some folks do not have any symptoms or signs at all [17].

1.1.3. Breast Cancer Stages and Survival Rate. Cells in the conduit lining or a segment of the breast are strange, as indicated by the SEER Committee's stage framework (surveillance, epidemiology, and end results). Breast disease is bound to create in one or the two breasts. At this moment, the perseverance rate is 100%. Breast disease is a sort of harmful development that impacts the tissues of the breast. The disease is short of one inch in broadness [25–27]. This stage has a 95% to 98 percent chance of perseverance. It has something to do with breast tissues, also. Growth is more modest than two crawls in measurement. The illness is malignant growth and may grow to the assistant lymph hubs. The endurance rate is currently at 88 percent breast tissues are affected. The tumor is more than two creeps in breadth. The disease might expand to the assistant lymph hubs. These things could occur as irritation, dimpling, or an adjustment of skin cooler. Right now, the endurance rate is somewhere close to half and 60%. Notwithstanding the breast, the disease has spread to different pieces of the body. Right now, the opportunity for endurance ranges somewhere in the range of 15%– 20%.

1.1.4. Breast Cancer Causes. The progression or change in the DNA of the cells is the reason for this. A percentage of the danger factors are benign conditions, such as

hyperplasia, which raise the risk of breast cancer. Having a history of malignant development increases your chances of creating disease [6].

1.2. Breast Cancer: Types, Statistics and Tools for Early Detection

1.2.1. Breast Cancer Types and Statistics. Breast malignant growth creates when cells in the breast start to multiply wildly. These cells generally combine to make cancer, which should be visible on an x-beam or as a felt knot. If the cells of the growth penetrate (attack) surrounding tissues or spread (metastasize) to other parts of the body, the cancer is dangerous (carcinogenic). Breast disease fundamentally influences ladies; however, it can likewise affect men. Breast disease can start in various areas of the Breast. By far, most breast tumors start in the pipes that convey milk to the areola (malignant ductal growths). A few malignancies (lobular growths) begin in the organs that make breast milk [14, 28].

It is basic to comprehend that by far, most breast knots are not threatening; all things being equal, they are harmless. Harmless breast cancers are harmless developments that do not spread outside of the breast and do not imperil one's life. Some harmless breast knots, then again, can build a lady's gamble of creating breast disease. A medical services expert ought to assess any bump or change in the breast to decide if it is harmless or dangerous and whether it might adjust future disease risk. Breast disease can possibly spread to different pieces of the body through the lymphatic framework [15].

The lymph framework is made up of y lymph centers, lymph chambers, and lymph liquids that can be tracked down all through the body [29–31]. Lymph centers are little bean-molded groups of resistant framework cells connected together by lymph (or lymphatic) veins. Lymph vessels are like little veins in that they transport lymph (as opposed to blood) out from the breast. Lymph contains both tissue liquids and waste, as well as insusceptible compromised framework cells. Breast disease cells are fit for entering lymphatic channels and topping off lymph hubs [18].

There is a chance that cancerous growth cells have spread to your lymph nodes, and a huge likelihood they have spread to different spots of your body (metastasized). Early location is basic for human existence, particularly on account of forceful breast malignant growth. Breast malignant growth is the most regular sickness among ladies around the world, as indicated by the World *Cancer* Research Fund International (WCRFI), with over 1.7 million new cases recognized in 2012. It is the second most normal disease on the planet [20, 32–34]. These record for around 12% of all new disease cases and 25% of all malignant growth cases in ladies.

Age standardized rate per 1, 00,000 (world).

2. Review of the Literature

Many studies on breast cancer in the field of medical data analysis have been distributed, and most of them guarantee to have extraordinary grouping exactness.

An assessment report has been distributed in which they characterized breast disease threatening development involving numerous techniques, including choice, highlight extraction, and component gathering strategies. The WEKA procedure is used to break down breast malignant growth illness data from the UCI store, and the portrayed strategies are utilized to arrange details exactly. This study demonstrates unequivocally that the information-gathering technique is efficient in predicting breast cancer sickness. The WEKA device is highly considered in data mining as one of the most effective and dependable data classification strategies. SVM outperformed other methods on the data set for breast cancer malignant development and produces consistent results. With an accuracy rate of 82.53 percent, 286 instances and 10 characteristics of breast cancer disease were investigated.

The authors of [6]fostered a framework plan for foreseeing the event of breast disease threatening development in the beginning phases by assessing the littlest assortment of factors from the clinical dataset. The Wisconsin breast malignant growth dataset (WBCD) was used to carry out the suggested investigation. The capacity of the proposed technique is determined by comparing actual with expected features and obtaining characterisation exactness. The results suggest that this inquiry has the highest categorization accuracy of 99.28%.

"Addresses the execution of a progress improvement algorithm to mine two distinctive datasets in AI using four unique progression smoothing techniques." Iris Data We used the set and the breast cancer dataset as informational data. A set to test the selected suggested improvement estimates [15].

The cerebrum affiliation (NN) is used in the solicitation issue of this exploration, close by four streamlined upgrade strategies: whale optimization algorithm (WOA), dragonfly algorithm (DA), multiverse optimization (MVA), gray wolf optimization (GWO). For precise choices of the proposed enhancement processes, an assortment of control limits was analyzed. Concerning intermingling, runtime, characterisation rate, and MSE, the comparison study shows that the GWO and MVO outperform both the WO and the DA. Hybrid algorithms that combine two different optimization methodologies could be investigated further in the future for data mining tasks" [20].

Recognized five managed AI computations are support vector machine (SVM), K-nearest neighbor method, random forest, artificial neural network (ANN), and logistic regression (LR) [1]. The UCI Storage House contains associations with the Wisconsin Breast Cancer Observation Dataset. Accuracy, mindfulness, specificity, accuracy, negative predictable values, false negative rates, interesting positive rates, $F1$ scores, and Matthews correlation coefficients are fully used to examine experimental results. The accuracy and accuracy of the ANN is only 98.57%, 97.82%, and 0.9890, which are the highest, while the accuracy and accuracy of the SVM are the second highest, with $F1$ scores of 97.14%, 97, 14%, and .9890, respectively.

"Working on the performance of most grouping algorithms by applying techniques for highlight choice to limit

how much elements" [20]. Certain properties are more important than others and affect the results of order calculations. The inspiration that drives this study is an overview of the accuracy of some recent information mining estimates in predicting the recurrence of breast disease. It utilizes molecule swarm advancement as a component choice in three notable classifiers: Naive Bayes, K-closest neighbor, and quick choice tree student to move along the prediction model's accuracy. The performance of naive Bayes was superior with and without PSO (particle swarm optimization), but when combined with PSO, the other two approaches improved."

Support vector machine (SVM)-based BCD model for distinguishing breast disease with 10-overlap cross approval. At the point when there are numerous information highlights for malignant growth ID, the undertaking gets more troublesome [15]. Principal component analysis (PCA) is a procedure for restricting the element region extending from a higher to a lower aspect. The PCA refines the model's accuracy based on the assessment results. Other learning calculations such as decision tree (DT), random forest, k-nearest neighbor method (k-NN), stochastic gradient descent (SGD), AdaBoost, neural network (NN), and naave Bayes have been proposed. It is displayed in a different way. BCD model. (query)According to the $F1$ measure, ROC twist, accuracy, lift curve, and calibration chart, the proposed BCD model provides excellent accuracy among the various model strategies investigated. The proposed BCD model has the most elevated AUC of 0.995 and the most noteworthy precision of 98.1 percent among different models tried."

The effectiveness of the five nonlinear machine learning methods is A.A. Evaluated in an 18-inch study by Bataineh: Multilayer Perceptron (MLP), K-nearest neighbor method (KNN), classification and regression tree (CART), Gaussian Naive Bayes (NB), and support vector machine (SVM) [37, 38] when it comes to data classification, the main goal is to evaluate the performance and feasibility of each algorithm in terms of accuracy, accuracy, and validation of characteristic tests. MLP has 96.70% accuracy when compiling data. This is more obvious than the other four algorithms. After evaluation, the accuracy and validation of the predictive model are evaluated by using the K-overlap mutual recognition method for ambiguous data. According to the results of this review, the accuracy, accuracy, and validation of the MLP model were highest at 99.12 percent, 99.00 percent, and 99.00 percent, respectively. The review relied on the Wisconsin Breast Cancer Diagnosis (WBCD) dataset.

Jing Zheng, Denan Lin, ZhongjunGao, Shuang Wang, Mingjie He, Jipeng Fan, Jing Zheng, Denan Lin Den Mathematically proposed an efficient AdaBoost algorithm for deep learning support to identify harmful enhancements to the breast (DLA-EABA) was run by using the current PC-development carried out. Notwithstanding customary PC vision methodologies, deep convolutional neural networks are being used to investigate tumour classification methods that use transfers (CNNs). This examination centers around deciding the best strategy by consolidating a few methodologies for choosing and separating highlights, as well as assessing their result utilizing grouping and division algorithms, which are essential for AI techniques. The

preliminary results exhibit a serious level of precision of 97.2 percent, responsiveness of 98.3 percent, and particularity of 96.5 percent when contrasted with other contemporary systems."

3. Research Methodology

Pertinent distributions for this audit were found via looking through Google Scholar, Research Gate, PubMed, Science Direct, IEEE, and Springer for terms like ["Breast *Cancer*" or "Dangerous development Detection"] + "Man-made intelligence" + ["SVM or ANN"] + [Biosensors or FET or Electrochemical"] [36, 37]. This audit takes care of every one of the applications where breast malignant growth recognition can be applied. Prior to 2018, studies were deemed less relevant because the approaches offered were less accurate, were more expensive, and had a smaller scope of implementation.

4. Data Analysis

Zheng, Jing, Lin, Denan, Gao, Zhongjun, Wang, Shuang, He, Mingjie und Fan, Jipeng Entwicklungen *Cancer* arrangement techniques that include moves, notwithstanding conventional PC vision strategies, are as of now being investigated utilizing profound convolutional brain organizations (CNNs). This exploration centers around deciding the ideal strategy by consolidating a few AI procedures with techniques for choosing and removing highlights, as well as assessing their result utilizing order and division calculations. When contrasted with other existing frameworks, the trial results show that the high exactness level of 97.2 percent, awareness of 98.3 percent, and particularity of 96.5 percent are superior.

4.1. A Comparative Analysis of Breast Cancer Prediction Techniques (Year 2016 to 2020). Further study gives a point-by-point portrayal of breast malignant growth grouping utilizing different AI, information mining, and profound learning procedures in light of the Algorithm/technique utilized for expectation, instruments, informational index, information type, and number qualities considered for the review, as displayed in Table 1.

5. Result Analysis

We searched the information databases of BMC Bioinformatics, Biomed, Google Scholar, IEEE, Science-Direct, Springer, and Web of Science for the penetration rate and research of huge learning systems in the fields of information mining, AI, and clinical information evaluation. Multiview-Mammography-based informative placement/numerical property-based informative combinations were found to have been used in the evaluation study [39–41]. This genuine Figure 3 and close to assessment Table 1 further shows that mammograms/pictures are used in relatively few examinations of breast-threatening development assumptions. The accessibility of datasets is a critical obstacle to utilizing AI and profound learning strategies to foresee breast malignant growth on the grounds that every strategy requires a lot of preparing

TABLE 1: Comparison of breast cancer data mining, machine learning, predictive deep learning techniques [19].

S. no	Author name	Method	Tools	Dataset	Data type	No. of attribute	Performance
1.	Venkateshwara Rao, Mary Gladence	Classification techniques SVM, Naïve Bayes	Weka	UCI machine learning	Numeric attributes	296	Higher accuracy 85%
2.	Madhu kumara, Vijhendra Singh	Classification supervised machine learning algorithm	MAtlab	Wisconsin breast cancer	Numeric	700	KNN classifier with 100%
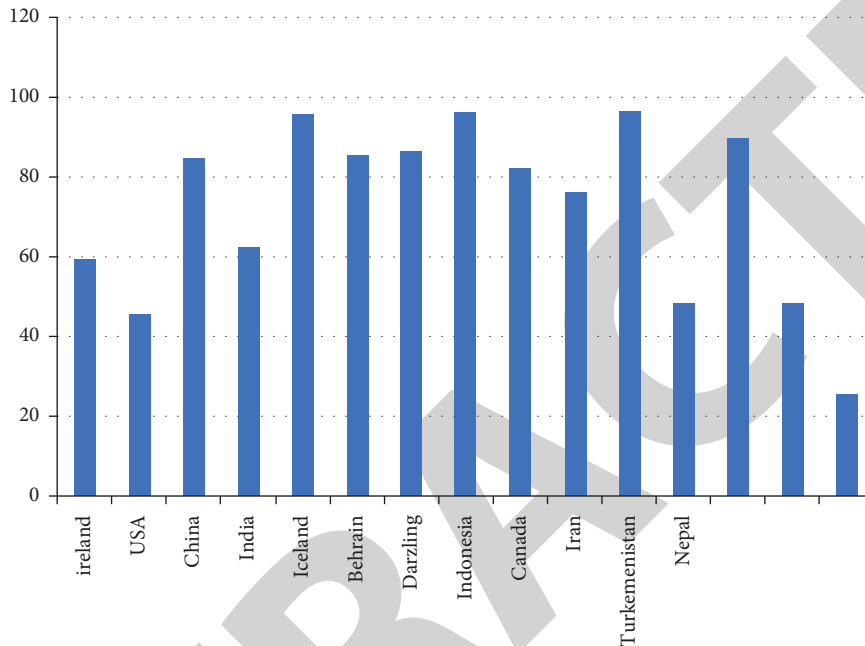
FIGURE 3: Statistical comparison of data mining, machine learning, and deep learning methods for breast cancer prediction over the last 5 years (2016–2020).

information for computational evaluations. We present an outline of information mining, AI, and profound learning approaches in this work, with an accentuation on the precision speed of breast infection assumption. We searched the information databases of BMC Bioinformatics, Biomed, Google Scholar, IEEE, Science-Direct, Springer, and Web of Science for the penetration rate and research of huge learning systems in the fields of information mining, AI, and clinical information evaluation. Multiview-Mammography-based informative placement/numerical property-based informative combinations were found to have been used in the evaluation study [37].

6. Conclusion

The basic role of this research article is to identify existing AI and important learning studies that assume the development of harmful breasts and identify the best framework for assessing their proportion. It has been perceived that more work still needs to be carried out from here on out since large information is creating an unrest in medical care at this moment [42, 43].

There is a need to deal with this enormous measure of information since the present advanced medical services require astute incorporation and collection of accessible patient data and PC information, coordinated, semi-

organized, and unstructured, in their unique organizations. Second, because of the shortage of accessible datasets, not very many exploration studies on breast disease pictures are led. Accordingly, a model for foreseeing breast malignant growth from a histopathological picture put together informational collections with respect to big information can be proposed. To start, Hadoop engineering can be made to hold information tests to imagine the work on Big information, and afterward, an effective convolutional brain network calculation can be developed for expectation [18].

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.

Acknowledgments

References

[1] M. Islam, M. d. RezwanulHaque, HasibIqbal, M. d. MunirulHasan, and M. N MahmudulHasan, "Breast Cancer Prediction: A comparative study using machine learning techniques," *SN COMPUTERS SCIENCE*, vol. 1290 pages, 2020.

[2] S. Eltalhi and H. Kutrani, "Breast cancer diagnosis and prediction using machine learning and data mining techniques: a review," *IOSR J. Dental Med. Science*, vol. 14, pp. 85–94, 2019.

[3] M. D. Ganggayah, N. A. Taib, Y. C. Har, P. Lio, and S. K. Dhillon, "'Predicting factors for survival of breast cancer patients using machine learning techniques," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, p. 48, 2019.

[4] A. A. Ibrahim, A. I. Hashad, and N. E. M. Shawky, "'A comparison of open source data mining tools for breast cancer classification," *Handbook of Research on Machine Learning Innovations and Trends*, pp. 636–651, IGI Global, Hershey PA USA, 2017.

[5] M. Hosni, I. Abnane, A. Idri, J. M. C. de Gea, and J. L. FernándezAlemán, "Reviewing Ensemble Classification Methods in Breast Cancer," *Computer Methods Programs Biomed*, vol. 177, pp. 89–112, Aug. 2019.

[6] MadhuKumari and s. Vijendra, "breast cancer prediction system" international conference on computational intelligence and data science," *Procedia Computer Science*, vol. 132, pp. 371–376, 2018.

[7] M. Abdar and V. Makarenkov, "CWV-BANN-SVM Ensemble Learning Classifier for an Accurate Diagnosis of Breast Cancer," *Measurement*, vol. 146, pp. 557–570, 2019.

[8] S. P. Rajamohana, A. Dharani, P. Anushree, B. Santhiya, and K. Umamaheswari, "Machine learning techniques for healthcare applications: early autism detection using ensemble approach and breast cancer prediction using SMO and IBK," *in Cognitive Social Mining Applications in Data Analytics and Forensics*, pp. 236–251, IGI Global, Hershey, PA, USA, 2019.

[9] M. Togacar and B. Ergen, "Deep learning approach for classification of breast cancer," in *Proceedings of the 2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, Malatya, Turkey, September 2018.

[10] M. Tiwari, R. Bharuka, P. Shah, and R. Lokare, *'Breast Cancer Prediction Using Deep Learning and Machine Learning Techniques,* " SSRN, New York NY USA, 2020.

[11] D. Selvathi and A. A. Poornila, "Deep learning techniques for breast cancer detection using medical image analysis," *in Biologically Rationalized Computing Techniques for Image Processing Applications*, pp. 159–186, Springer, Cham, Switzerland, 2018.

[12] M. Badr, S. Al-Otaibi, N. Alturki, and T. Abir, "Detection of heart arrhythmia on electrocardiogram using artificial neural networks," *Computational Intelligence and Neuroscience*, pp. 1–10, 2022.

[13] A. Hiba, h. HajarMousannifb, c. Al moatassime, and t. noeld, "Using machine learning algorithms for breast cancer risk prediction and diagnosis," in *Proceedings of the 2018 3rd International Conference on Circuits, Control, Communication and Computing (I4C)*, Bangalore India, October 2018.

[14] Alghunaim and H. H. Al-Baity, "On the Scalability of Machine-Learning Algorithms for Breast Cancer Prediction in Big Data Context," *IEEE Access*, pp. 91535–91546, 2019.

[15] A. M. Hemeida a, b Salem Alkhalaf, A. Mady c, E. A. Mahmoud, M. E. Hussein, and M. Ayman, d. BahaEldin, Implementation of nature-inspired optimization algorithms in some data mining tasks" Published by Elsevier B.V. on behalf of Faculty of Engineering," *Ain Shams University. Ain Shams Engineering Journal*, vol. 11, no. 2, pp. 309–318, June 2020.

[16] L. VenkateswaraRao, V. Mary Gladence and R. Lakshmi, Research of Feature Selection Methods to Predict Breast Cancer," *International Journal of Recent Technology and Engineering*, vol. 17, pp. 2356–2367, 2019.

[17] H. Dhahri, E. Al Maghayreh, A. Mahmood, W. Elkilani, and M. Faisal Nagi, "Automated breast cancer diagnosis based on machine learning algorithms," *Journal of Healthcare Engineering*, vol. 11, 2019.

[18] C. Walid, "Optimization of K-NN Algorithm by Clustering and Reliability Coefficients: Application to Breast-Cancer Diagnosis," *Elsevier Procedia Computer Science*, vol. 127, pp. 293–299, 2018.

[19] H. Huang, Xi'an Feng, S. Zhou et al., "A new fruit fly optimization algorithm enhanced support vector machine for diagnosis of breast cancer based on high-level features," *BMC Bioinformatics*, vol. 20, no. 8, p. 290, 2019.

[20] N. B. SapiahBintiSakri, A. Rashid, and Z. Muhammad, "Particle Swarm Optimization Feature Selection for Breast Cancer Recurrence Prediction," *Special Section on Big Data Learning and Discovery IEEE Access*, vol. 6, pp. 29637–29647, 2018.

[21] G. Hamed, M. A. E.-R. Marey, S. E.-S. Amin, and M. F. Tolba, "Deep learning in breast cancer detection and classification," in *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020)*, Cham Switzerland, March 2020.

[22] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajjej, and A. A. Shihab, "Applying dynamic systems to social media by using controlling stability," *Computational Intelligence and Neuroscience*, pp. 1–7, 2022.

[23] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[24] S. Khalil, L. Hatch, C. R. Price et al., "Addressing Breast Cancer Screening Disparities Among Uninsured and Insured Patients: A Student-Run Free Clinic Initiative," *J Community Health*, vol. 45, no. 15, pp. 1–5, 2019.

[25] A. Memis, N. Ozdemir, M. Parildar, E. E. Ustun, and Y. Erhan, "'Mucinous (colloid) breast cancer: mammographic and US features with histologic correlation," *European Journal of Radiology*, vol. 35, no. 1, pp. 39–43, 2000.

[26] A. Reddy, B. Soni, and S. Reddy, "Breast Cancer Detection by Leveraging Machine Learning," *ICT Express*, vol. 6, no. 4, pp. 320–324, 2020.

[27] Z. Salod and Y. Singh, "'Comparison of the performance of machine learning algorithms in breast cancer screening and detection: a protocol," *J. Public Health Res.* vol. 8, no. 3, p. 1677, Dec. 2019.

[28] M. Supriya and A. J. Deepa, "A Novel Approach for Breast Cancer Prediction Using Optimized ANN Classifier Based on Big Data Environment" Health Care Management Science," *Springer, Science+Business Media, LLC, part of Springer Nature*, vol. 23, no. 10, pp. 414–426, 2019.

[29] C. Siotos, A. Naska, R. J. Bello et al., "'Survival and Disease Recurrence Rates Among Breast Cancer Patients Following Mastectomy with or without Breast Reconstruction," *Plastic Reconstructive Surg*, vol. 144, no. 2, pp. 169e–177e, 2019.

Research Article

Efficacy of Yishen Huashi Granules Combined with Linagliptin Tablets on Blood Glucose and Renal Function in Patients with Type 2 Diabetic Nephropathy

Panke Zhang [ID],[1] Jingxi Meng,[2] Mingliang Duan,[1] Dan Li,[1] and Ruixin Wang[1]

[1]*Department of Nephrology Rheumatology, Zhengzhou Hospital of Traditional Chinese Medicine, Zhengzhou 450000, Henan, China*
[2]*Department of Diabetes, Zhengzhou Hospital of Traditional Chinese Medicine, Zhengzhou 450000, Henan, China*

Correspondence should be addressed to Panke Zhang; zhangpanke@zzszyy.org.cn

Objective. To probe into the efficacy of Yishen Huashi granules combined with linagliptin tablets in the treatment of type 2 diabetic nephropathy (DN) and its effect on blood glucose and renal function in patients. *Methods*. 70 patients with type 2 DN at our hospital between May 2020 and May 2022 were chosen as the research objects and separated into the control group and the research group based on their treatments. With 35 cases in each group, the patients treated with initial therapy and linagliptin tablets were enrolled in the control group, and those who received the above treatments and also Yishen Huashi granules were included in the research group. Their clinical indexes such as blood glucose and renal function were compared with both groups after treatment. *Results*. After treatment, the research group had remarkably lower fasting blood glucose (FPG), 2 h-postprandial blood glucose (2 h-PBG), and glycosylated hemoglobin A1c (HbA1c) levels than those in the control group ($P < 0.05$). After treatment, the research group had remarkably lower levels of total cholesterol (TC), triglyceride (TG), and low-density lipoprotein (LDL) ($P < 0.05$) and higher high-density lipoprotein (HDL) levels ($P < 0.05$) than those in the control group. After treatment, the urinary microalbumin (u-mALB) level was remarkably lower in both groups ($P < 0.05$) and was distinctly lower in the research group than in the control group ($P < 0.05$). After treatment, the research group had remarkably lower renal function indexes such as serum creatinine (SCr), blood urea nitrogen (BUN), urinary protein (UPro), and urinary albumin excretion rate (UAER) ($P < 0.05$) and a higher estimated glomerular filtration rate (eGFR) level ($P < 0.05$) than those in the control group. The efficacy was evaluated by the traditional Chinese medicine (TCM) syndrome score after treatment. There were no patients in complete remission between both the groups, where slight differences were found in the proportion of significant remission ($P > 0.05$), with the total effective rate of the research group remarkably higher than that of the control group ($P < 0.05$). *Conclusion*. The combination of Yishen Huashi granules and linagliptin tablets can reduce the blood glucose and blood lipid levels in patients with type 2 DN and lower UPro and protect renal function at the same time, which provides a new idea and a method for clinical treatment of type 2 DN with integrated traditional Chinese and Western medicine.

1. Introduction

Diabetic nephropathy (DN), characterized by persistent proteinuria, decreased glomerular filtration rate (GFR), and elevated blood pressure, is the major primary disease of end-stage nephropathy and one of the lethal microvascular complications of diabetes, with an incidence rate reaching up to 30%–50% [1–3]. DN has an insidious onset, which progresses rapidly by the proteinuria stage in the clinic, with symptoms such as edema, dyslipidemia, and continuous decline in renal function. However, conventional hypoglycemic agents are limited in DN treatment, and most oral hypoglycemic agents cannot be used in stages 3–4 of chronic kidney disease (CKD) [4–6]. In contrast, as a new type of a dipeptidyl peptidase 4 (DPP-4) inhibitor, linagliptin tablets can be excreted through the intestine and can be used even in

patients with renal insufficiency or on dialysis, with a good hypoglycemic effect. In addition, traditional Chinese medicine (TCM) treatment has been applied to DN and CKD for more than 2,000 years, whose herbal medicine is still widely adopted in treating DN conditions. From the perspective of TCM, DN is a syndrome of qi and yin deficiency with blood stasis, whose treatment should be based on nourishing qi and yin, activating blood circulation, and removing blood stasis. Yishen Huashi granules, which are composed of ginseng, astragalus, large head atractylodes rhizome, Poria cocos, Rhizoma Alismatis, ternate pinellia rhizome, notopterygium root, radix angelicae tuhuo, divaricate saposhnikovia root, radix bupleuri, etc., have the effect of invigorating yang and spleen, tonifying kidneys, dissipating hygrosis, and inducing diuresis and detumescence. Pharmacological studies have also confirmed that the granules can reduce UPro, regulate immune function, inhibit oxidation, reduce glycemic indexes, and repair damaged glomerular basement membranes [7–9]. So far, there has been no research combining these two in the treatment of DN. In our hospital, we have been using Yishen Huashi granules in combination with linagliptin tablets to treat type 2 DN for many years with good efficacy, which is reported as follows.

2. Materials and Methods

2.1. Inclusion Criteria. ① The patients' disease met the clinical diagnostic criteria of type 2 DN [10] and was stable; ② the patients were aged 18 and older; ③ the patients' gender was not limited; ④ the patients had good compliance with treatment and follow-ups; ⑤ the patients did not take any vitamins, statins, or angiotensin-converting enzyme inhibitors within the last 30 days; ⑥ the patients' disease met the indications for Yishen Huashi granules and linagliptin tablets;⑦ the patients and their family members knew the purpose and procedure of this study and signed the informed consent.

2.2. Exclusion Criteria. ① Patients with severe liver and kidney dysfunction, malignant tumor, or coagulation disorders; ② patients receiving maintenance dialysis or renal transplantation; ③ patients with other primary or secondary nephropathy; ④ patients with diabetic ketoacidosis; ⑤ patients with hyperthyroidism or severe infection;⑥ patients with renovascular hypertension.

2.3. Selection and Grouping of Patients. 70 patients with type 2 DN at our hospital between May 2020 and May 2022 were chosen as research objects to conduct a retrospective analysis. They were separated into a control group and a research group based on the treatment methods, with 35 cases in each group. This study conformed to ethical standards and was authorized by the ethics committee of our hospital.

2.4. Treatment Methods. Patients were all given conventional treatment of DN, including high-quality-low-protein,

low-phosphorus diabetic diet, control of blood glucose and blood pressure, correction of anemia, correction of acidosis, correction of disorders of calcium and phosphorus metabolism, regulation of blood lipids, maintenance of electrolyte balance, and prevention and control of infection [11–13]. Based on this, the control group took linagliptin tablets (specification: 5 mg; manufacturer: Boehringer Ingelheim Pharmaceuticals Inc.; NMPA Approval No. J20171087) orally, 5 mg per day, which could be taken at any time of the day, with or without meals. With the above treatments, the research group took Yishen Huashi granules (specification: 10 g/bag; manufacturer: Guangzhou Consun Pharmaceutical Co., Ltd.; NMPA Approval No. Z20090250), 1 bag each time, 3 times a day. The observation treatment period of both groups was 3 months.

2.5. Observation Criteria

(1) General data include age, gender, duration of diabetes, concomitant risk factors, body mass index (BMI), fasting blood glucose (FPG), 2 h-postprandial blood glucose (2 h-PBG), glycosylated hemoglobin A1c (HbA1c), total cholesterol (TC), triglyceride (TG), serum creatinine (SCr), blood urea nitrogen (BUN), tumor necrosis factor-α (TNF-α), interleukin-6 (IL-6), high-sensitivity C-reactive protein (hs-CRP). SCr is generally considered to be endogenous SCr, while endogenous creatinine is a product of human muscle metabolism. BUN is a nitrogenous compound in plasma other than protein, which is excreted from the body by glomerular filtration. BUN will increase when renal insufficiency is decompensated. The presence of protein in the urine is called proteinuria or UPro. Normal urine contains a small amount of small molecule protein, and proteinuria is a common manifestation of kidney disease and can also occur in systemic diseases.

(2) Glycemic indexes. The glycemic indexes of patients in both groups were measured after treatment, including FPG, 2h-PBG, and HbA1c indexes.

(3) Blood lipid levels. The blood lipid levels in the two groups were tested after treatment, including TC, TG, high-density lipoprotein (HDL), and low-density lipoprotein (LDL) indexes.

(4) Urinary microalbumin (u-mALB). The u-mALB levels were tested before and after treatment in patients.

(5) Renal function includes SCr, BUN, estimated GFR (eGFR), urinary protein (UPro), and urinary albumin excretion rate (UAER) indexes.

(6) Efficacy in treating DN assessed by the TCM syndrome score. Patients who had negative results of proteinuria, UPro quantification of <0.5 g/24h, normal or near normal serum albumin levels (>35 g/L), normal or near normal urinary erythrocyte levels, normal renal function, and complete elimination of clinical nephropathic symptoms by multiple

detections were complete remission. Patients who had UPro quantification of <1 g/24h, remarkably improved serum albumin and urinary erythrocyte levels, and normal or near normal renal function by multiple detections were significant remission. Multiple detection results of decreased UPro levels, UPro quantification reduced by half compared with that before treatment, improved serum albumin and urinary erythrocyte levels, and better renal function were partial remission. UPro, serum albumin, and urinary erythrocyte levels had no remarkable changes compared with those before treatment, with obvious clinical nephropathic symptoms and unchanged renal function as invalid treatment.

2.5.1. Laboratory Testing. Venous blood samples were taken from patients before and after treatment for laboratory testing, with the indexes of FPG, HbA1c, TC, TG, SCr, BUN, HDL, LDL, TNF-α, IL-6, and hs-CRP. 24-h urine samples were collected before and after treatment, 50 ml of which were taken for testing, with the indexes of u-mALB, eGFR, UPro, and UAER.

2.6. Statistical Disposal. The study adopted SPSS22.0 for data processing, which mainly calculated the differences between the groups, and graph production was carried out using GraphPad Prism 7 (GraphPad Software, San Diego, USA). The research data contained two types of data, count data and measurement data, in which the former was represented as [n (%)] and verified by X^2 and the latter was represented as ($\overline{x} \pm s$) and verified by t tests, conformed to normal distribution. The statistical result of $P < 0.05$ indicated a statistical difference between the groups.

3. Results

3.1. General Data. No statistical differences were found in the age, genders, duration of diabetes, concomitant risk factors, BMI, FPG, 2h-PBG, HbA1c, TC, TG, SCr, BUN, TNF-α, IL-6, and hs-CRP between both the groups ($P > 0.05$). The general data in both groups were balanced and comparable, see Table 1.

3.2. Glycemic Indexes. The glycemic indexes of all the patients showed reductions after treatment compared to those before treatment, with the FPG, 2 h-PBG, and HbA1c levels remarkably lower in the research group than those in the control group ($P < 0.05$), see Table 2.

3.3. Blood Lipid Levels. After treatment, the research group had remarkably lower TC, TG, and LDL levels ($P < 0.05$) and higher HDL levels ($P < 0.05$) than those in the control group. The blood lipid levels in the research group after treatment were eminently better than those in the control group, see Table 3.

3.4. U-mALB Levels. After treatment, the u-mALB level was remarkably lower in both groups ($P < 0.05$) and was distinctly lower in the research group than in the control group ($P < 0.05$), see Figure 1.

3.5. Renal Function. After treatment, the research group had remarkably lower renal function indexes such as SCr, BUN, UPro, and UAER ($P < 0.05$) and a higher eGFR index ($P < 0.05$) than those in the control group, see Table 4.

3.6. Efficacy of TCM Syndrome. The efficacy was evaluated by the TCM syndrome score after treatment. There were no patients in complete remission between both the groups, where slight differences were found in the proportion of significant remission ($P > 0.05$), with the total effective rate of the research group being remarkably higher than that of the control group ($P < 0.05$), see Figure 2.

4. Discussion

Type 2 DN is a microvascular complication often taking place in diabetic patients, with insidious onset, less obvious initial symptoms, and malignant changes in renal function, when the disease develops to the late stage. Diabetes is the major factor of end-stage renal disease, and a clinical study has confirmed [14] that DN occurs in most patients 15–20 years after diabetes, which is one of the lethal hazards to diabetic patients. However, in recent years, TCM treatment has been widely carried out in clinics, bringing more opportunities and possibilities for treating type 2 DN. TCM believes that DN belongs to the categories of "edema" and "guange" (frequent vomiting and dysuria), which is developed on the basis of the deficiency of qi and yin in diabetes and obstruction of collaterals by blood stasis and toxin. The pathogenesis is mainly the deficiency of the spleen and the kidney, incompetence of qi transformation, and internal resistance of blood stasis [15, 16]. Due to long-term diabetes and weak qi of the spleen, the source of qi and blood is blocked, resulting in the kidney's failure of storing energy and dysfunction of water metabolism. Ginseng, large head atractylodes rhizome, and Poria cocos in Yishen Huashi granules have the effect of invigorating yang and qi and strengthening the spleen to eliminate dampness, which can remove the influence of damp in middle jiao on the kidney. *Pinellia ternata* rhizome, rhizoma alismatis, and divaricate saposhnikovia root enable the recovery of renal function by eliminating dampness and swelling and tonifying qi and the kidney. Rhizoma coptidis, the root of herbaceous peony, and dried old orange peel can remove dampness. For DN, there are limitations in the treatment of Western medicine or TCM alone. Most conventional oral hypoglycemic agents in the clinic need to be excreted by the kidneys, and thus, most patients need to receive insulin therapy. But as renal clearance of insulin decreases and the duration of insulin action prolongs, the risk of hypoglycemia gradually increases, and even cardiovascular and cerebrovascular events can be induced [17, 18]. Linagliptin is a xanthine derivative which has the advantages of strong activity, high DPP-4

TABLE 1: Comparison of general data ($n = 35$).

Observation indexes	Control group	Research group	X^2/t	P
Age (years)	58.51 ± 8.30	58.34 ± 8.40	0.085	0.932
Genders			0.238	0.626
Male	20 (57.14)	22 (62.86)		
Female	15 (42.86)	13 (37.14)		
Duration of diabetes (years)	5.89 ± 1.12	6.00 ± 1.15	0.405	0.687
FPG (mmol/L)	10.25 ± 0.67	10.34 ± 0.71	0.545	0.587
2h-PBG (mmol/L)	12.56 ± 1.97	12.50 ± 1.88	0.130	0.897
HbA1c (%)	8.80 ± 1.47	8.78 ± 1.50	0.056	0.955
TC (mmol/L)	5.47 ± 0.65	5.48 ± 0.53	0.071	0.944
TG (mmol/L)	2.08 ± 1.40	2.13 ± 0.38	0.204	0.839
SCr (umol/L)	73.05 ± 16.85	72.96 ± 17.10	0.22	0.982
BUN (mmol/L)	8.53 ± 5.59	8.47 ± 5.41	0.046	0.964
Concomitant risk factors				
Hypertension	29 (82.86)	31 (88.57)	0.467	0.495
Hyperlipidemia	20 (57.14)	19 (54.29)	0.058	0.810
BMI (kg/m^2)	20.86 ± 2.01	21.05 ± 2.10	0.387	0.700
TNF-α (ng/L)	141.84 ± 24.10	142.48 ± 23.71	0.112	0.911
IL-6 (ng/L)	60.07 ± 8.37	60.20 ± 8.85	0.063	0.950
Hs-CRP (mg/L)	7.55 ± 1.85	7.48 ± 1.83	0.159	0.874

TABLE 2: Comparison of glycemic indexes after treatment.

Groups	Number of cases	FPG (mmol/L)	2 h-PBG (mmol/L)	HbA1C (%)
Control group	35	8.70 ± 0.90	11.07 ± 1.97	8.55 ± 0.70
Research group	35	7.33 ± 0.80	10.10 ± 1.65	7.46 ± 0.68
t		6.731	2.233	6.608
P		<0.001	0.029	<0.001

TABLE 3: Comparison of blood lipid levels after treatment ($n = 35$).

Groups	TC (mmol/L)	TG (mmol/L)	HDL (mmol/L)	LDL (mmol/L)
Control group	4.70 ± 0.27	2.02 ± 0.26	1.03 ± 0.08	3.18 ± 0.29
Research group	4.23 ± 0.24	1.85 ± 0.22	1.16 ± 0.13	2.51 ± 0.16
t	6.697	2.953	5.038	11.968
P	<0.001	0.004	<0.001	<0.001
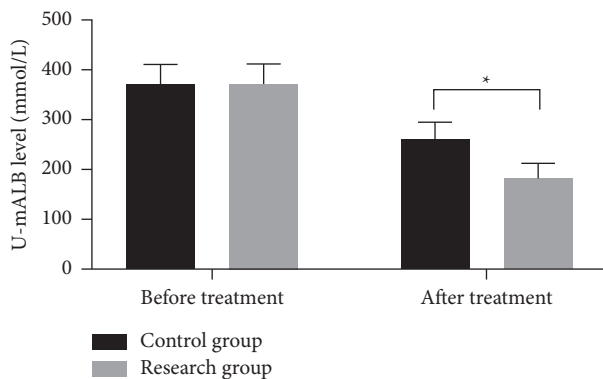
FIGURE 1: Comparison of u-mALB levels. Notes: the transverse axis was before and after treatment, with the longitudinal axis as the u-mALB level (mmol/L). The u-mALB levels in the control group before and after treatment were (370.37 ± 40.36) and (260.78 ± 34.24), respectively, while in the research group were (371.45 ± 40.58) and (182.13 ± 30.56). * indicated a remarkable difference in the u-mALB levels between both the groups after treatment ($t = 10.138$, $P < 0.001$).

receptor selectivity, long half-life, and a high protein binding rate. Most of the drug is excreted via the intestinal tract as a prototype, with only 5% by the kidney. It is the only DPP-4 inhibitor that does not require dose adjustment in the treatment of DN, which has good hypoglycemic effects even in CKD stage 5 and can continue to be used in the presence of hepatic insufficiency [19, 20]. The combined application of Linagliptin and Yishen Huashi granules may create synergy, which is beneficial to improve the clinical efficacy of DN patients.

4.1. Effect on Glycemic Indexes and Blood Lipid Levels.
The results revealed that all the patients' glycemic indexes showed different degrees of reduction after treatment compared to those before treatment, with the FPG, 2hPG, and HbA1C levels in the research group remarkably lower than those in the control group ($P < 0.05$). After treatment, the research group had remarkably lower TC, TG, and LDL levels ($P < 0.05$) and higher HDL levels ($P < 0.05$) than those in the control group, suggesting eminently better glycemic

TABLE 4: Comparison of renal function indexes after treatment ($n = 35$).

Groups	SCr (umol/L)	eGFR (ml/min·1.73 m^2)	BUN (mmol/L)	UPro (g/24h)	UAER (ug/min)
Control group	67.33 ± 8.14	80.64 ± 6.46	7.21 ± 2.15	1.15 ± 0.22	76.15 ± 10.72
Research group	62.12 ± 6.93	86.12 ± 10.96	6.23 ± 1.55	1.00 ± 0.18	70.83 ± 8.34
t	2.883	2.548	2.187	3.122	2.317
P	0.005	0.013	0.032	0.03	0.024
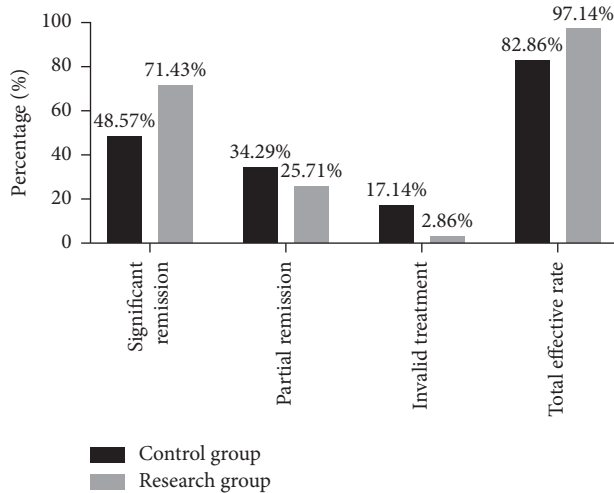
FIGURE 2: Comparison of efficacy of TCM syndrome after treatment ($n = 35$). Notes: the transverse axis represented evaluation dimensions, with the longitudinal axis as the percentage (%). After treatment, there were 17 cases of significant remission, 12 cases of partial remission, 6 cases of invalid treatment, and a total of 29 cases of effective treatment in the control group; there were 25 significant remissions, 9 partial remissions, 1 invalid treatment, and a total of 34 effective treatment in the research group. * suggested a remarkable difference in the total effective rate of treatment between both the groups ($X^2 = 3.968$, $P = 0.046$).

indexes and blood lipid levels in the research group after treatment than those in the control group. Previous studies have shown [21–23] that Yishen Huashi granules can control blood glucose and regulate blood lipids. From the standpoint of TCM, hyperlipidemia belongs to the category of "stagnation of damp turbidity and static blood," which is mainly treated by invigorating yang and discharging turbidity, activating blood circulation, and removing blood stasis. Yishen Huashi granules have the functions of invigorating yang and the spleen, eliminating dampness and turbidity, and inducing diuresis and detumescence. According to modern pharmacology, astragalus can improve blood microcirculation, and rhizoma alismatis can inhibit the occurrence of hyperglycemia and hypersensitivity, thus stabilizing blood glucose and blood lipids and reducing the occurrence of glomerular basement membrane lesions. Therefore, Yishen Huashi granules can effectively alleviate the symptoms and signs of edema, fatigue, inappetence, chillness, and cold limbs caused by DN.

4.2. Effect on Renal Function. According to the analysis on the renal function-related indexes of the two groups, the u-mALB level was remarkably lower in both groups after treatment ($P < 0.05$) and was distinctly lower in the research group than in the control group ($P < 0.05$). After treatment, the research group had remarkably lower renal function indexes such as SCr, BUN, UPro, and UAER ($P < 0.05$) and a higher eGFR index ($P < 0.05$) than those in the control group, suggesting that the combination of Yishen Huashi granules and linagliptin had a better effect on improving renal function and reducing renal injury. Modern pharmacology has shown that ginsenosides in Yisheng Huashi granules can reduce proteinuria, diminish inflammation and water retention, and protect renal function; polysaccharides in Poria cocos activate the cellular immune response and improve patients' low immunity, thereby alleviating the inflammatory response and promoting the repair of renal tubular damage [24, 25]. With the addition of linagliptin tablets, this combination can exert renal protective effects while lowering blood glucose. Moreover, the efficacy evaluated by the TCM syndrome score after treatment showed that there were no patients in complete remission between both the groups, where slight differences were found in the proportion of significant remission ($P > 0.05$), with the total effective rate of the research group remarkably higher than that of the control group ($P < 0.05$). It has further confirmed that Yishen Huashi granules combined with linagliptin tablets have a good effect of glucose control on patients with type 2 DN and contributes to the reduction of renal injury, with a definite curative effect and further promoting the rehabilitation of patients.

5. Summary

The combination of Yishen Huashi granules and linagliptin tablets can effectively reduce blood lipids and control blood glucose, with the effect of lowering UPro and protecting renal function, which provides a new idea and a method for clinical treatment of type 2 DN with integrated traditional Chinese and Western medicine. It is speculated in the study that lipid lowering may have an obvious effect on controlling UPro, improving hypercoagulability, and delaying the disease progression of DN patients. But it is worth noting that the sample size selected for this study was limited and that the findings may be influenced by factors such as geographical factors, so the correlation can be confirmed by the forward large-sample studies.

Data Availability

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

Conflicts of Interest

The authors declare that they not have no conflicts of interest.

Acknowledgments

References

[1] J. Zhang, J. Li, and J. Q. Huang, "Network meta-analysis of four Chinese patent medicines combined with angiotensin converting enzyme inhibitors or angiotensin receptor blockers in early diabetic nephropathy treatment," *World J Tradit Chin Med*, vol. 6, no. 1, pp. 51–60, 2020.

[2] X. P. Yan, W. R. Wang, H. Y. Liu et al., "Effect of bushen qingre yuyin decoction on salivary secretion, spleen index, submandibular gland index, submandibular gland histomorphology, and aqp5 expression in the nonobese diabetic mouse model," *World J Tradit Chin Med*, vol. 6, no. 4, pp. 508–514, 2020.

[3] A. J. Scheen, "Pharmacokinetic characteristics and clinical efficacy of an SGLT2 inhibitor plus DPP-4 inhibitor combination therapy in type 2 diabetes," *Clinical Pharmacokinetics*, vol. 56, no. 7, pp. 703–718, 2017.

[4] B. A. Moussa, M. A. Mahrouse, and G. M. Fawzy, "Application of experimental design in HPLC method optimization and robustness for the simultaneous determination of canagliflozin, empagliflozin, linagliptin, and metformin in tablet," *Biomedical Chromatography: Biomedical Chromatography*, vol. 35, no. 10, p. e5155, 2021.

[5] G. Goderis, B. Vaes, P. Mamouris, E. van Craeyveld, and C. Mathieu, "Prevalence of atherosclerotic cardiovascular disease, Heart failure, and chronic kidney disease in patients with type 2 diabetes mellitus: a primary care research network-based Study," *Experimental and Clinical Endocrinology and Diabetes*, vol. 130, no. 7, pp. 447–453, 2022.

[6] A. R. Gosmanov, D. E. Gemoets, L. S. Kaminsky, C. P. Kovesdy, and E. O. Gosmanova, "Efficacy of metformin monotherapy in US veterans with type 2 diabetes and pre-existing chronic kidney disease stage 3," *Diabetes Obesity and Metabolism*, vol. 23, no. 8, pp. 1879–1885, 2021.

[7] Y. Tu and G. L. Lichuanping, "Efficacy and safety analysis of Yishen Huashi granule in the adjuvant treatment of early diabetes nephropathy," *World Journal of traditional Chinese medicine*, vol. 15, no. 7, pp. 1042–1046, 2020.

[8] Y. Zhang, "Clinical study of Yishen Huashi granule combined with valsartan in the treatment of chronic glomerulonephritis," *World Journal of Traditional Chinese Medicine*, vol. 12, no. 1, pp. 68–70, 2017.

[9] L. WeiLei, "Analysis of 63 cases of purpura nephritis treated with Yishen Huashi granule and prednisone," *Anhui medicine*, vol. 24, no. 1, pp. 166–169, 2020.

[10] S. Mahmud and A. Koratala, " Assessment of venous congestion by doppler ultrasound: a valuable bedside diagnostic tool for the new-age nephrologist," *CEN Case Reports*, vol. 10, no. 1, pp. 153–155, 2021.

[11] L. He, H. Wang, C. Gu, X. He, L. Zhao, and X. Tong, "Administration of traditional Chinese blood circulation activating drugs for microvascular complications in patients with type 2 diabetes mellitus," *Journal of Diabetes Research*, vol. 2016, no. 6, pp. 1081657–1081665, 2016.

[12] L. Xiang, P. Jiang, L. Zhou et al., "Additive effect of qidan dihuang grain, a traditional Chinese medicine, and angiotensin receptor blockers on albuminuria levels in patients with diabetic nephropathy: a randomized, parallel-controlled trial," *Evidence-based Complementary and Alternative Medicine*, vol. 2016, no. 5, Article ID 1064924, 8 page, 2016.

[13] S. Ono, Y. Ono, D. Koide, and H. Yasunaga, "Association between routine nephropathy monitoring and subsequent change in estimated glomerular filtration rate in patients with diabetes mellitus: a Japanese non-elderly cohort study," *Journal of Epidemiology*, vol. 30, no. 8, pp. 326–331, 2020.

[14] Q. Kong, H. Zhang, T. Zhao et al., "Tangshen formula attenuates hepatic steatosis by inhibiting hepatic lipogenesis and augmenting fatty acid oxidation in db/db mice," *International Journal of Molecular Medicine*, vol. 38, no. 6, pp. 1715–1726, 2016.

[15] Z. Yunyun, C. Wenjuan, S. Lan, and W. Hai, "Fecal metabonomic study of a polysaccharide, MDG-1 from Ophiopogon japonicus on diabetic mice based on gas chromatography/time-of-flight mass spectrometry (GC TOF/MS)," *Molecular BioSystems*, vol. 10, no. 2, pp. 304–312, 2014.

[16] L. U. Wang, C. Yang, F. Song, Z. Liu, and S. Liu, "Therapeutic effectiveness of *Gardenia jasminoides* on type 2 diabetic rats: mass spectrometry-based metabolomics approach," *Journal of Agricultural and Food Chemistry*, vol. 68, no. 36, pp. 9673–9682, 2020.

[17] H. Wang, Y. Teng, S. Li et al., "UHPLC-MS-Based serum and urine metabolomics reveals the anti-diabetic mechanism of ginsenoside Re in type 2 diabetic rats," *Molecules*, vol. 26, no. 21, p. 6657, 2021.

[18] A. B. Moussa, A. M. Marianne, and M. G. Fawzy, "A validated LC-MS/MS method for simultaneous determination of linagliptin and metformin in spiked human plasma coupled with solid phase extraction: application to a pharmacokinetic study in healthy volunteers," *Journal of Pharmaceutical and Biomedical Analysis: An International Journal on All Drug-Related Topics in Pharmaceutical, Biomedical and Clinical Analysis*, vol. 163, no. 30, pp. 163153–163161, 2019.

[19] R. Meza-Palacios, A. A. Aguilar-Lasserre, E. L. Urena-Bogarin, C. F. Vázquez-Rodríguez, R. Posada-Gómez, and T. A. Armín, "Development of a fuzzy expert system for the nephropathy control assessment in patients with type 2 aiabetes mellitus," *Expert Systems with Application*, vol. 72, pp. 335–343, 2017.

[20] C. Q. F. Klessens, M. Zandbergen, R. Wolterbeek et al., "Macrophages in diabetic nephropathy in patients with type 2 diabetes," *Nephrology Dialysis Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, vol. 32, no. 8, pp. 1322–1329, 2017.

[21] L. Zhao, Y. Zhang, F. Liu et al., "Urinary complement proteins and risk of end-stage renal disease: quantitative urinary proteomics in patients with type 2 diabetes and biopsy-proven diabetic nephropathy," *Journal of Endocrinological Investigation*, vol. 44, no. 12, pp. 2709–2723, 2021.

[22] S. Zhang, X. Li, H. Luo, Z. Z. Fang, and H. Ai, "Role of aromatic amino acids in pathogeneses of diabetic nephropathy in Chinese patients with type 2 diabetes," *Journal of Diabetes and its Complications*, vol. 34, no. 10, Article ID 107667, 2020.

[23] H. Nacaixia, "Zhao Jing Effects of gligliptin on glucose and lipid metabolism and islets of Langerhans in type 2 diabetes patients with obesity β Effect of cell secretory function," *Chinese Journal of clinical practical medicine*, vol. 11, no. 2, pp. 37–40, 2020.

[24] C. Zhao, J. Hu, Z. Wang, and Z. Y. Cao, "Serum LncRNA PANDAR may act as a novel serum biomarker of diabetic nephropathy in patients with type 2 diabetes," *Clinical Laboratory*, vol. 66, no. 6, pp. 1067–1072, 2020.

[25] L. Ildiko, B. Nadine, S. Regina et al., "Triple fixed-dose combination empagliflozin, linagliptin, and metformin for patients with type 2 diabetes," *Postgraduate Medicine*, vol. 132, no. 4, pp. 337–345, 2020.

Retraction

Retracted: Mobility and Quality of Life in Chinese Elderly and Geriatric Patients and Biomedical Diagnosis

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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Xin and L. Li, "Mobility and Quality of Life in Chinese Elderly and Geriatric Patients and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7457161, 9 pages, 2022.

Research Article

Mobility and Quality of Life in Chinese Elderly and Geriatric Patients and Biomedical Diagnosis

Jing Xin [1,2] **and Liyuan Li**[2]

^{1}School of Public Administration, Zhongnan University of Economics and Law, Wuhan 430073, China
^{2}Fanling Business School, Nanyang Institute of Technology, Nanyang 473004, China

Correspondence should be addressed to Jing Xin; 19404245@masu.edu.cn

Mobilization hugely affects the personal satisfaction of old people. We found that preparing old and geriatric patients expands their personal satisfaction using statistical comparisons of observed groups. As a technique of evaluation, we employed the nonparametric Mann–Whitney test. The Pearson correlation coefficient was employed to determine relationships between variables. In both the BREF questionnaire (item 15) and the ADL test, geriatric individuals had similar mobility rates. Physical activity is a crucial feature of mobility in seniors, according to findings from other studies. In hospitalized geriatric patients, reduced mobility is the most significant impediment. Seniors' confidence in their own talents grew as a result of the favourable overall influence of physical activity. Geriatric patients must be mobilized on a regular basis in nursing, depending on their level of dependency. Physical activity had a positive overall effect on older people's confidence in their own abilities, and this was shown in how they felt about themselves. Nursing staff must constantly move geriatric patients based on how much assistance they require.

1. Introduction

Actual work limitations, long haul disease, and advanced age may not generally mean a lower personal satisfaction. More seasoned individuals can get the things that help them, similar to social combination, idealism, certainty, and a longing to carry on with a full life, from the things that help them get them. In practice, this means that the medical findings of elderly patients may not always correspond to their current psychological and social status. This is because people who are older have bad moods, a lack of purpose in life, a lot of addictions (depending on others), and less control over their own lives [1].

Patient fulfilment is one of the principle signs of how great a medical care administration is. It is essential to zero in on addressing the requirements of patients, particularly as they connect with their real diseases, in the way that nurses care today. How many patients are satisfied with a certain medical device is an important way to figure out how good it really is. The quality of health care and how satisfied patients

are with their care are important indicators for caregivers and health facility managers. Our goal was to find out how important mobility is to the personal satisfaction of geriatric patients and individuals who live in care offices for the older. We imagined that the portability of geriatric patients and individuals who live in care offices for the old is a significant piece of how they rate their personal satisfaction and that it would be very different between groups [2].

1.1. Geriatric Patients.
There is a field of medicine called geriatrics that deals with medical care for older people, but it is not easy to figure out what that age group is. A lot of people use the word "older" instead of "elderly." People often use the age of 65, but most people do not need geriatric care until they are 70, 75, or even 80. What does gerontology look at? Gerontology is the study of ageing. It looks at changes in biology, sociology, and psychology as we get older.

Around 1900, in the United States, people over 65 made up 4% of the population. Today, they make up more than

14% (nearly 50 million people, with a daily net gain of 10,000). More than 80 million people will be over 65 by 2026, when baby boomers born after World War II reach age 80. Those who are over 65 now have an average age of a little more than 75. The proportion of those over 85 is expected to grow fastest, though [3].

Men have 17 more years at 65, and women have 10 more years at 75. Ladies typically live about 5 years longer than men. This is most likely due to genetic, natural, and environmental factors. Women's lifestyles changed dramatically in the late twentieth and early twenty-first centuries, with more smoking and stress. These differences in survival have not changed, though as show in Figure 1 [4].

1.2. Purposes and Components of Comprehensive Geriatric Assessment.

The evaluation attempts to accurately diagnose, screen for treatable illnesses, devise a reasonable treatment plan, and track any changes in elderly people over time. Often, the appropriateness of services (like long-term care) and the best place to put them are also looked into. In CGA, a lot of different people work together to solve problems. Multidisciplinary teams can add a lot of expertise and enthusiasm to the way patients are assessed and cared for [5]. The final structure will depend on the goals of the programme, the setting, the number of patients, and the financing. Most gatherings have a specialist, a medical caretaker, and a social labourer. Others in the group are a physiotherapist and a word-related advisor. A couple of gatherings similarly have a dietitian, advisor, trained professional, podiatrist, ophthalmologist, or clinical pharmacologist on them, too. Since there are so many various things that can turn out badly with individuals as they age, the WHO has distinguished and suggested that specific things be checked in old individuals (Box). In view of these areas, various instruments have been used to evaluate people [6].

Geriatric instruments are usually scales that have been proven to be accurate and easy to use. Another benefit is that standardized instruments can make it easier for healthcare workers to share information. This allows for smooth teamwork, the collection of valuable and true data, and the tracking of treatment progress over time [7].

People use the Barthel ADL and Katz index of ADL to determine how simple a person's daily living is. To determine the instrumental ADL, tools such as the Lawton IADL scales and the Fillenbaum five-item IADL questionnaire are used (IADL). Folstein Mini-Mental State Examination (FMSSE) is two tests that measure how well people think that FMSE is an example of scales that can be used to measure cognitive function (MMSE) [8].

There is proof that these tests work [9].

It is also possible to use a lot of different tools to screen for depression. When they want to know how depressed they are, many people use the Yesavage Geriatric Depression Scale, the Hamilton Rating Scale for Depression, Self-Care (D), and the Zung Self-Rating Depression Scale.
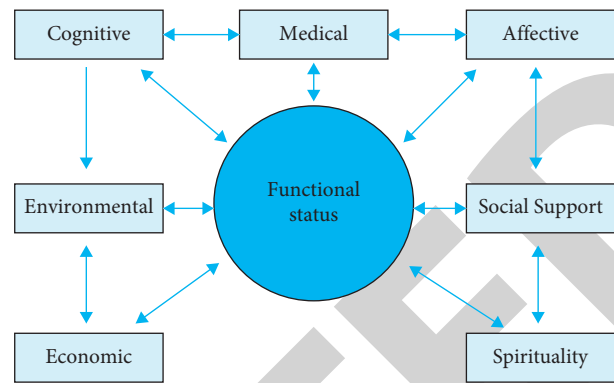
FIGURE 1: Geriatric assessment.

People in the area have used a Cantonese-language version of the Geriatric Depression Scale. It has been shown to be true. The only way to measure social and economic status does not seem to be with a single number. Any such measure would be very long because it would have to cover a lot of ground look at a lot of different things about how socially important you are. When the Royal College of Physicians of London recommends a checklist, this is a simple one [10].

The most difficult aspect of a geriatric assessment is determining how physically fit someone is. This is mostly because physical illness is so complicated and has so many different causes. There are some well-designed scales available, but they are only useful for certain diseases, such as the New York Heart Association's classification for heart disease, which is widely used [11].

As a result, many CGA programmers do not try to figure out how many physical illnesses their participants have, but only record their diagnoses and the medicines they use. One of the most common things that people do not pay attention to when they do a physical examination on an older person is to check for ailing health. Nourishing issues are normal in more established patients, particularly the individuals who have other medical conditions. Henceforth, a nourishing appraisal ought to be important for the actual assessment in the CGA, which is the reason [12–14].

The body weight, weight file, mid-arm outline, rear arm muscles skin overlay thickness, and serum egg whites level can be used to figure out how well an elderly person is getting the nutrients they need. It can be hard to measure the height of an elderly person because they have problems with their bones and are not able to move. You can use the length of their arms to figure out their height [8].

1.3. Benefits of Using Comprehensive Geriatric Assessment.

Many distinct and controlled studies have been conducted to determine how well CGA functions. Probably, the latest investigations that show the advantages of CGA are displayed in Table 1 and Figure 2. Old individuals who were probably going to be shipped off a nursing home yet were

TABLE 1: Those who did controlled studies were able to show the benefits of the comprehensive geriatric assessment.

Study	Patient setting	Year	Benefit found*							
			1	2	3	4	5	6	7	8
Retention at*	Inpatient	19481986								
Landed et al.*	Inpatient	1970	*	*			*		*	
Honan et al.*	Inpatient	1970	*	*		*				
Bout et al.*	Inpatient	1963	*	*		*		*		
Silver man et al.*	Outpatient	1970	*	*					*	
Vitter et al.*	Outpatient	1955	*				*			
Headship et al.*	Inpatient	1970	*		*				*	

not excessively wiped out, hysterical, or well enough to be sent to a geriatric evaluation unit were randomized to go there (intervention group). There were geriatricians, a social worker, and geriatric nurses on staff at the unit. It was also helped by occupational and physiotherapists, psychologists, and dietitians. Compared with the control group, those who took part in the study had a lower mortality rate, fewer people who went to nursing homes, a shorter stay in the hospital, less money spent on institutional care, and better morale and function [15, 16].

The accuracy of tests has been improved. 6 mean less nursing home use. The placement of 2 changed for the better. There has been a lot more use of home healthcare service. The function has been made better by number 3, less use of medical service by number 8, and a better mood or thought process by number 4 [17–20].

Older people who were placed in a geriatric-focused special unit had higher ADL scores, fewer nursing home admissions, and overall better well-being and health than those who were not have done two controlled studies in people. 34,35 In the 1987 study, the older people who were checked out by the geriatric consultation service but did not get any follow-up care had better mental health, took fewer medications when they left the hospital, and died less quickly than the people who did not get checked out. 34 In a later study,35 the people who took part in the study were given follow-up [21, 22]. It was found that 6-month survival rates went up, as well as a better Barthel index and less use of institutional care 35 [5].

In the same way, researchers have said that CGA can help people who are not in the hospital. Doing a controlled study, Boult and his team36 looked at 43 elderly people who had a high chance of being readmitted. They took them to an outpatient geriatric assessment and the executives follow up for 3 to 4 months, where a geriatric multidisciplinary group dealt with them. Following 17 months, there was a lower passing rate and less utilization of nursing homes and crisis administrations. The consequences of another review utilized short-term geriatric evaluation without recovery benefits or direct control of the executives [23, 24]. They had the option to show that the mediation bunch had a superior analysis, better mental and passionate result, and a lower level of carer stress than the benchmark group. Other community-based controlled trials have found that CGA is beneficial to people 38, 39 [8].

2. Materials and Method

We used ex-post facto causal comparative study in our investigation. We employed descriptive statistics to confirm the findings of the study techniques. Our study included 386 patients from the geriatric departments of Presov's College Hospital with J. A. Reiman Policlinic and Bardejov's Hospital and Clinic of St. James, as well as offices for the elderly in nearby towns. We utilized the methodology of gathering information utilizing a normalized poll, the WHOQOL-BREF survey, and the everyday exercises of day-to-day living, ADL test, to decide the singular parts of exploration [25–29]. The two studies were completed entirely by each patient/senior, and we took into account the unique characteristics of senior age, particularly tactile hindrance (changes in vision, hearing, and so on). The average length of stay in a senior facility was 13.8 months (SD 1.29). We utilized the Mann–Whitney test for two independent selections to compare the observed groups statistically. We were able to find statistically significant differences in measured parameters across groups using this nonparametric methodology. The nonparametric Spearman correlation coefficient was employed to determine relationships between variables [6].

2.1. Research Methodology. This is how we did our work: we used causal comparative analysis, which is done after the fact. To make sure that our research methods worked, we used descriptive statistics. As part of our study, we looked at more than 400 patients in geriatric departments at the University Hospital in Presov and Bardejov, as well as at facilities for older people in both cities and at other places where people can go to get help. To figure out which parts of the research were important, we utilized the technique for getting information from a normalized survey, the WHO-QOL-BREF poll, and the everyday exercises of day-to-day living, ADL test [30–33]. The two polls were finished up with every patient/senior all alone, and we considered the extraordinary parts of senior age, like tangible hindrance, while finishing them up (changes in vision, hearing, and so on). Consistently, the normal length of stay in a consideration home for the older was 13.8 months (SD 1.29). For factual examinations of the gatherings we saw, we utilized the Mann–Whitney test. This nonparametric technique has permitted us to track down genuinely huge contrasts between bunches in the boundaries that were measured. We used the nonparametric, Pearson correlation coefficient to figure out how the two variables were related [5].

2.2. Data Analysis

2.2.1. Descriptive Statistics. An example of a descriptive stat is a summary stat that quantitatively describes or sums up features from a group of data. Descriptive statistics are the process of using and analysing these stats (see Table 2 and Figure 3).

There is no doubt that patients in care offices for the old are better ready to move (thing 20). Nonetheless, they show a
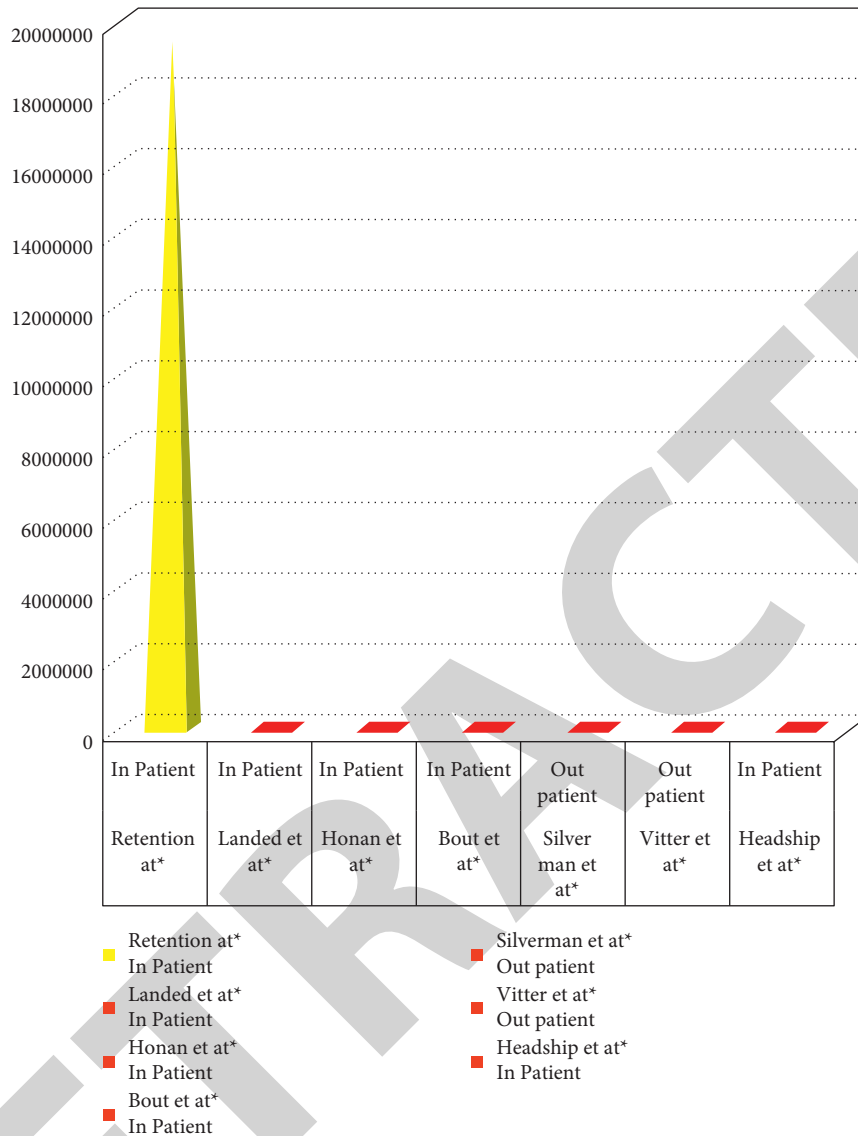
Figure 2: Graphical representation of table.

Table 2: Statistical results of differences investigated between the two groups identified by Mann–Whitney test.

Variable	Geriatric patients		Facilities for seniors		z
	M	SD	M	SD	
BREF 15	3.65	1.45	2.22	2.17	−3.625***
ADL moving	3.56	1.36	4.33	2.48	−2.532**

Significance level: *$_{p<0.05}$，**$_{p<0.01}$，**$_{p<0.001}$.

lower pace of versatility in the ADL test since they utilize more assistive gadgets while strolling (supports and wheelchairs). We think this is on the grounds that this gathering has a great deal of long haul medical conditions. They have shown a similar degree of versatility in the BREF poll and the ADL test. This demonstrates how important mobility and self-sufficiency are to these patients (Table 1 and Figure 1) [7].

2.2.2. Correlation. We can see that geriatric patients who are more mobile in the questionnaire WHOQOL-BREF (item 20) are if they are more mobile, they have a higher likelihood of having a good quality of life and a lower likelihood of having a good quality of life. The same thing happened to a group of people who lived in nursing homes of elderly people. In both of the groups we looked at, more mobility means that the overall value of quality of life goes up (Tables 3 and 4, Figures 4 and 5) [10].

3. Result and Discussion

Physical exercise is an important factor in the mobility of older people, according to research from around the world. The standard range SWLS looked at the quality of life around the world (satisfaction with life scale). He discovered that age ($p = 0.05$) was related to physical activity (=0.34), own performance (=0.30), personal values (=0.22), and satisfaction (=0.12) using descriptive statistics. According to him,
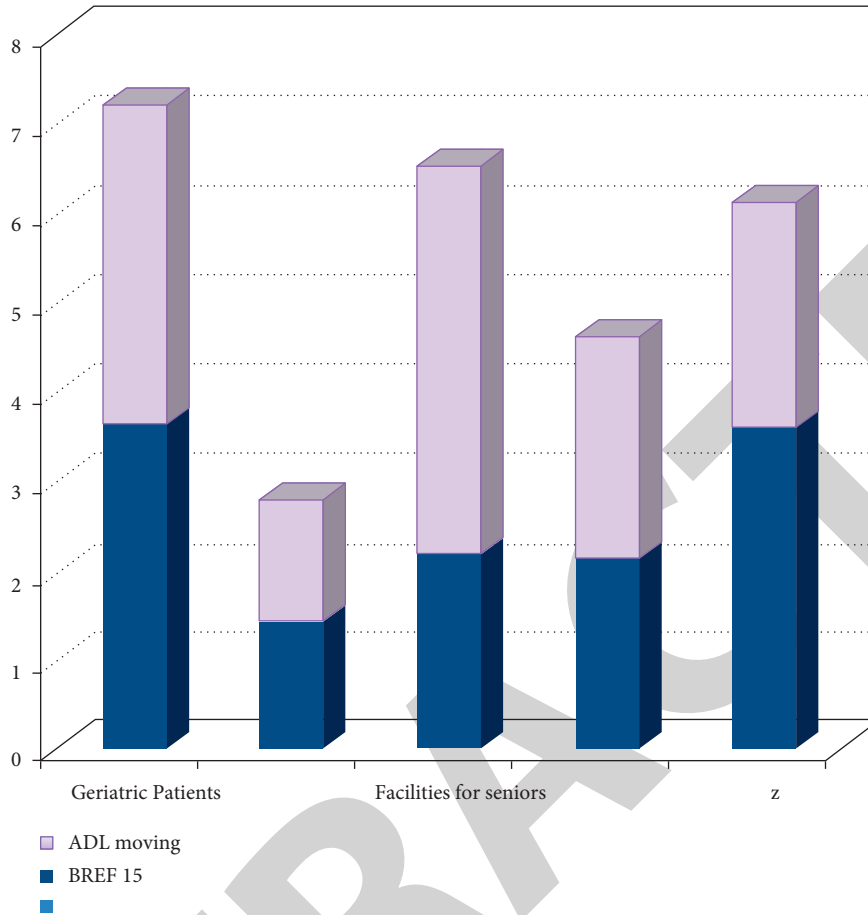
FIGURE 3: Comparison of mean values using WHOQOL-BREF questionnaire and using questionnaire in geriatric patients.

TABLE 3: Statistical results of Pearson correlation among the survey items.

Variable	BREF 20	BREF gross score	ADL mobility
BREF 20	—	0.472***	−0.754***
BREF gross score	0.452***	—	−0.325*
ADL mobility	−0.531***	−0.353***	—

Significance level: * $p < 0.05$, ** $p < 0.01$, ** $p < 0.001$.

TABLE 4: Statistical results of Pearson correlation among the survey items.

Variable	BREF 20	BREF gross score	ADL mobility
BREF 20	—	0.762***	−0.654***
BREF gross score	0.753***	—	−0.625***
ADL mobility	−0.631***	−0.453***	—

Significance level: * $p < 0.05$, ** $p < 0.01$, ** $p < 0.001$.

personal values play a significant role in the outcomes of physical exercise and can be influenced by factors such as happiness [34–36]. However, the fact is that the impact of demographic characteristics on the examined relationships is crucial. The author offers programmer to increase physical activity based on these findings. It also confirms the belief that seniors who engage in physical activity have a different perspective on their impairment status [3].

As a result, international studies reveal that this means that mobility is an important part of the quality of life for people who are old. In our study, we found that mobility is important for senior quality of life, and we found that it had an important level of significance of p (0.001). We found that elderly people in our study had about the same scores on the BREF scale (item 20), mobility ($M = 3.65$), and the ADL test ($M = 3.56$). A geriatric patient's nursing and physiotherapy teams work well together critical for these reasons. Setting up the rehabilitation is what we recommend.

With the use of the ADL test, mobility is revealed to be equal in both study groups (M = approximate 4). We recognize that physical incapacity entails the accumulation of difficulties that lead to immobility syndrome rather than post-treatment nursing, which primarily entails the management of senior patient health complications. The Faculty of Physical Education and Sport at The Charles University in Prague wrote a paper titled "Physical exercise as an active way of life for older people." They said that movement is the foundation of older people's quality of life. Its goal was to prove that a reasonable life in old age is possible. This is what mobility and relaxation activities that are good for older people are all about. People over the age of 60 were asked to help with the project through a monthly magazine that was published in Prague. It had more than 100 people come, which shows that the elderly are interested in physical
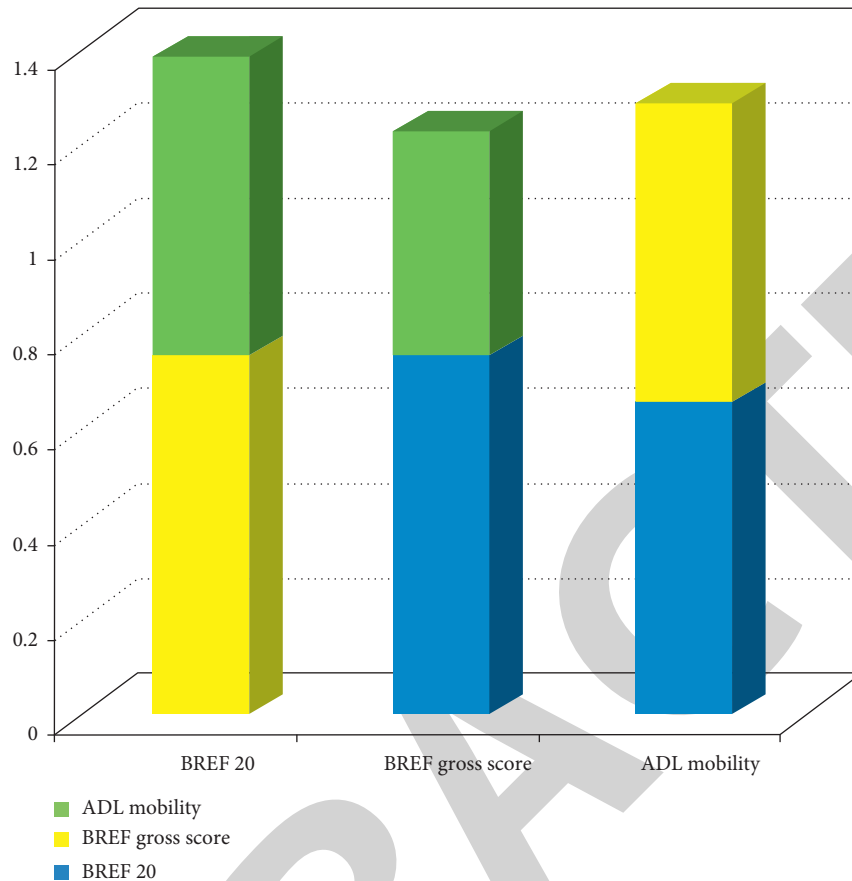
Figure 4: Statistical results of Pearson correlation in graphical form.

activity. Due to a lack of time, only 30 seniors were looked after. Some candidates were not chosen because they had a good doctor's report. The participants ranged in age from 62 to 78 years old. Throughout the project's 3.5-year run, the long-term influence of regular physical activity on one's lifestyle and sense of personal freedom was assessed. Throughout the experiment, participants completed a variety of functional, motor, and psychosocial tests, all of which were evaluated [3].

The results of the last set of tests and self-tests for older people show that delayed ageing has led to functional, psychological, and motor activity in older people. Physical activity has a positive effect on older people, such as when they have more self-confidence because of it. Another good thing about the programme was that it drew the attention of medical and social institutions that work with the elderly [15]. As people get older, their muscle mass goes down, and their bodies store fat and connective tissue as fat and muscle. In an 80-year-old person, muscle strength is down about 30%, and muscle fibres are lost by about 40% to 50%. Much research has shown that the average amount of muscle mass in seniors has gone down from 1.5 to 2 cm over the years. Their waist circumference has grown from 3 to 4 cm. They are more muscular, and their joints are in better shape [37–39]. People lose mobility in their joints and their ability to move their bodies in all different ways [8]. The value of how quickly your heart rate moves is the most common way to judge how well you did (hereafter SF). A scale called Karvonen's scale is used to measure the maximum value of SF. When SF max is reached, the heart rate is 220 beats per minute, which is the same as the scale. The person's age can be found at any time and without a lot of work. To check on the elderly who live in social homes, we can now do so. We want to start a group of senior nurses and physiotherapists who would run the group. To be a nurse, you would have to work very hard to keep an eye on the body's functions, which would show how important SF is, how healthy the senior is, and how motivated he is [3].
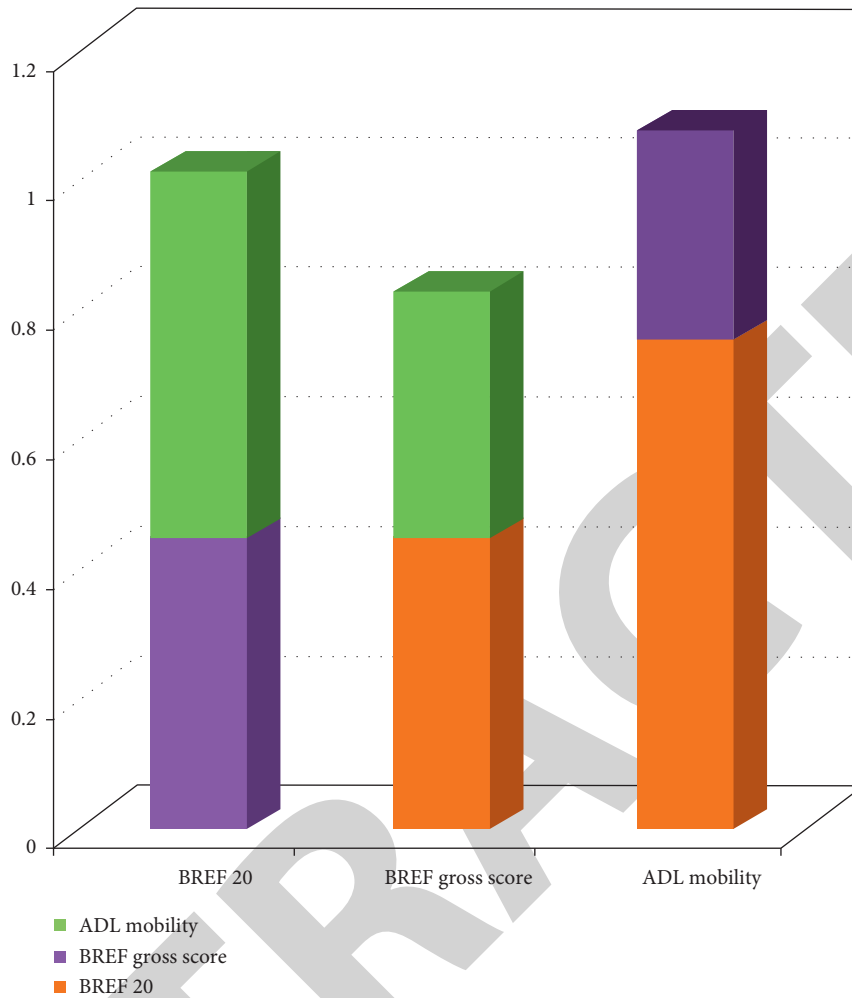
FIGURE 5: Statistical results of Pearson correlation.

4. Conclusions

The analysis of life in elderly population is motivated by demographic characteristics. The ageing population poses a problem to governments, especially in terms of health care. The average life expectancy at birth is increasing; men lived 70.85 years and 78.73 years. Reference [7, 8] said that Europe will get a lot of old people as the world gets older. By 2050, only one-third of Europe's people will be 60 or older [2]. In 2004, Canada signed a treaty that talked about how important it is to have a good quality of life for everyone, even people who are older. This is the goal of the national old programme: to keep and improve older people's self-sufficiency, social involvement, and integration as well as their general level of well-being. These recommendations and our research show that there is a real need to get older people moving and help them become more self-sufficient in social institutions. Geriatric age has many different characteristics, but from the standpoint of nursing, patient satisfaction in the hospital is a good way to tell if the care is good [22, 40]. Nursing staff must constantly move geriatric patients based on how much assistance they require. It is also important to work with physiotherapists to come up with training programmes for geriatric patients and to make sure they have enough tools to help them become more self-sufficient in their daily lives (their continuous use) [3].

Data Availability

Data that support the findings of this study are available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. No. 576/2004, *On Health Care, Services Related to Health Care and Amending Some Laws, the 243rd Amount § 2 Instruments: 12/01/10*, http://www.zbierka.sk/zz/predpisy/default.aspx?cc=243&ro=20%2004, 2004.

[2] B. IDVPZ, "Aging population. Instruments," p. 118, 2008, http://www.euractiv.sk/verzia-pre-tlac/zoznam_liniek/starnutieobyvatelstva.

[3] C. J. Brown, R. J. Friedkin, and S. K. Inouye, "Prevalence and outcomes of low mobility in hospitalized older patients,"

Journal of the American Geriatrics Society, vol. 52, no. 8, pp. 1263–1270, 2004.

[4] I. 20/01/10, "Collection of Laws Act 351/2005, which amends Act no. 578/2004 Z.z. on health care providers, health care professionals and organizations in health care," 2004, http://www.zbierka.sk/zz/predpisy/default.aspx?cc=243&ro=2004.

[5] K. Dunn, A. Torres, and J. Tiscani, "Functional status outcomes in a quality of life study with latinas," *Journal of Multicultural Nursing and Health*, vol. 10, p. 2, 2004.

[6] M. Eberhardt, D. Ingrfam, and D. Makuc, *Urban and Rural Health Chartbook. Health United States*, National Center for Health Statistics, Hyattsville, MD USA, 2001.

[7] L. Hegyi, Š Krajcik, and M. Final, "The current state of health care for seniors in SR," *Geriatric Times*, vol. 10, no. 3, pp. 1335–1850, 2009.

[8] Z. Kalvach, Z. Zadak, and R. Jirak, "Geriatrics and gerontology. 1," in *Motivational Elements when Working with Seniors*, D. Klevetova and I. Dlabalova, Eds., p. 864, Grada Publishing, Praha, Hlavni mesto Praha, 2004.

[9] E. Mcauley, J. F. Konopack, K. S. Morris et al., "Research priorities," *J. Geronte. Psychol. Sci. Soc. Sci.*vol. 61, pp. 267–273, 2006.

[10] K. Rosengren, "Physical activity and functional limitations in older women: influence of self-efficacy," *J. Geronte. Psychol. Sci. Soc. Sci.*vol. 60, pp. 901–909, 2005.

[11] J. K. Luk, K. H. Or, and J. Woo, "Using the comprehensive geriatric assessment technique to assess elderly patients," *Hong Kong medical journal = Xianggang yi xue za zhi*, vol. 6, no. 1, pp. 93–98, 2000, þ.

[12] N. J. Vetter, D. A. Jones, and C. R. Victor, "Effect of health visitors working with elderly patients in general practice: a randomised controlled trial," *BMJ*, vol. 288, no. 6414, pp. 369–372, 1984.

[13] C. Hendriksen, E. Lund, and E. Stromgard, "Consequences of assessment and intervention among elderly people: a three year randomised controlled trial," *BMJ*, vol. 289, no. 6457, pp. 1522–1524, 1984.

[14] A. M. Epstein, J. A. Hall, M. Fretwell et al., "Consultative geriatric assessment for ambulatory patients," *JAMA*, vol. 263, no. 4, pp. 538–544, 1990.

[15] M. Stilec, *The Active Lifestyle for Seniors*, p. 136, Jessica Kingsley, London, 2004.

[16] S. M. White, T. R. Wojcicki, and E. Mcauley, "Physical Activity and Quality of Life in Community Dwelling Older Adults," 2008, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2649048.

[17] T. A. Teasdale, L. Shuman, E. Snow, and R. J. Luchi, "A comparison of placement outcomes of geriatric cohorts receiving care in a geriatric assessment unit and on a sgeneral medicine floor," *J Am GeriatrSoc*, vol. 31, pp. 529–534, 1983.

[18] D. B. Reuben, G. M. Borok, G. Wolde-Tsadik et al., "A randomized trial of comprehensive geriatric assessment in the care of hospitalized patients," *New England Journal of Medicine*, vol. 332, no. 20, pp. 1345–1350, 1995.

[19] C. H. Winograd, M. B. Gerety, and N. A. Lai, "A negative trial of inpatient geriatric consultation. Lessons learned and recommendations for future research," *Archives of Internal Medicine*, vol. 153, no. 17, pp. 2017–2023, 1993.

[20] M. D. Fretwell, P. M. Raymond, S. T. McGarvey et al., "The senior care study," *Journal of the American Geriatrics Society*, vol. 38, no. 10, pp. 1073–1081, 1990.

[21] B. D. Hogan, R. A. Fox, B. W. Badley, and O. E. Mann, "Effect of a geriatric consultation service on management of patients in an acute care hospital," *Canadian Medical Association Journal*, vol. 136, pp. 713–717, 1987.

[22] D. B. Hogan and R. A. Fox, "A prospective controlled trial of a geriatric consultation team in an acute-care hospital," *Age and Ageing*, vol. 19, no. 2, pp. 107–113, 1990.

[23] M. Silverman, D. Musa, D. C. Martin, J. R. Lave, J. Adams, and E. M. Ricci, "Evaluation of outpatient geriatric assessment: a randomized multi-site trial," *Journal of the American Geriatrics Society*, vol. 43, no. 7, pp. 733–740, 1995.

[24] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajjej, and A. A. Shihab, "Applying dynamic systems to social media by using controlling stability," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–7, 2022, þ.

[25] A. E. Stuck, A. L. Siu, G. D. Wieland, L. Rubenstein, and J. Adams, "Comprehensive geriatric assessment: a meta-analysis of controlled trials," *The Lancet*, vol. 342, no. 8878, pp. 1032–1036, 1993.

[26] C. H. Winograd, M. B. Gerety, E. Brown, and V. Kolodny, "Targeting the hospitalized elderly for geriatric consultation," *Journal of the American Geriatrics Society*, vol. 36, no. 12, pp. 1113–1119, 1988.

[27] A. B. Lehmann, "Review: undernutrition in elderly people," *Age and Ageing*, vol. 18, no. 5, pp. 339–353, 1989.

[28] A. W. Mcevoy and O. F. W. James, "Anthropometric indices in normal elderly subjects," *Age and Ageing*, vol. 11, no. 2, pp. 97–100, 1982.

[29] M. B. Alazzam, A. T. Al-Radaideh, N. Binsaif, A. S. AlGhamdi, and M. A. Rahman, "Advanced deep learning human herpes virus 6 (HHV-6) molecular detection in understanding human infertility," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–5, 2022, þ, Article ID 1422963.

[30] T. Kwok and M. N. Whitelaw, "The use of armspan in nutritional assessment of the elderly," *Journal of the American Geriatrics Society*, vol. 39, no. 5, pp. 492–496, 1991.

[31] D. Podsiadlo and S. Richardson, "The timed "Up & Go": a test of basic functional mobility for frail elderly persons," *Journal of the American Geriatrics Society*, vol. 39, no. 2, pp. 142–148, 1991.

[32] D. K. Welner, P. W. Duncan, J. Chandler, and S. A. Studenski, "Functional reach: a marker of physical fraility," *J Am GeriatrSoc*, vol. 40, pp. 203–207, 1992.

[33] M. E. Tinetti, T. Franklin Williams, and R. Mayewski, "Fall risk index for elderly patients based on number of chronic disabilities," *The American Journal of Medicine*, vol. 80, no. 3, pp. 429–434, 1986.

[34] B. Robinson, "Validation of a caregiver strain index," *J Gerontol*, vol. 38, pp. 344–348, 1983.

[35] P. P. Vitaliano, J. Russo, H. M. Young, J. Becker, and R. D. Maiuro, "The screen for caregiver burden," *The Gerontologist*, vol. 31, no. 1, pp. 76–83, 1991.

[36] J. M. Kinney and M. A. P. Stephens, "Caregiving hassles scale: assessing the daily hassles of caring for a family member with dementia," *The Gerontologist*, vol. 29, no. 3, pp. 328–332, 1989.

Research Article

Big Data Analytics in Supply Chain Management: A Qualitative Study

Basim Aljabhan [1] **and Melese Abeyie** [2]

^{1}*Ports and Maritime Transportation Department, Faculty of Maritime Studies, King Abdulaziz University,*
 Jeddah 21589, Saudi Arabia
^{2}*College of Natural and Computational Science, Department of Chemistry, Injibara University, Injibara, Ethiopia*

Correspondence should be addressed to Melese Abeyie; meleseabiye@inu.edu.et

This work explores the leading supply chain processes impacted by big data analytic techniques. Although these concepts are being extensively applied to supply chain management, the number of works that examine and classify the main processes in the current literature is still scarce. This article, therefore, provides a classification of the current literature on the use of big data analytics and provides insight from professionals in the field in relation to this topic. A well-established set of practical guidelines was used to design and carry out a systematic literature mapping. A total of 50 primary studies were analysed and classified, chosen from a sample of 5, 437 studies after careful filtering to answer six research questions. In addition, a survey was prepared and applied by professionals working in the area. In total, 25 professionals answered a questionnaire with eleven questions, ten of which seek to explore the importance of big data analytics for the areas of the supply chain addressed in this work, and one intends to list the three areas where BDA can be more shocking. More than 60% of the studies are directly linked to the area of chain management; most studies performed empirical studies but rarely classified or detailed methodological procedures; almost 50% bring models to optimize some process or forecasts for better decision-making; more than 50% of professionals working in the area believe that the processes where big data analytics can effectively contribute are related to inventory and stockout management. This study serves as a basis for further research and future work, as it reviews the literature, pointing out the main areas that are being addressed and making a relationship with understanding these areas in practice.

1. Introduction

Technological breakthroughs are transforming data generation and analysis. According to Bumblauskas et al. [1], big data have the potential to alter management and the entire business process. This is why the transformation concept is so crucial. Kaynak et al. [2] define big data as a corporate resource with 5 Vs: volume, speed, variety, verifiability, and value. According to Kaur et al. [3], "volume" refers to the amount of data or physical space necessary to store it, which is expanding at an exponential rate, putting strain on existing storage systems. According to the same authors' definition of speed, how quickly data are transferred affects both data intake and production. Aye et al. [4] define a

variety of data as data that can be created from a range of platforms. This means there are no standards; each platform generates data based on its structure. As a result, semi-organized and fully organized data can be generated. Elgendy et al. [5] define veracity as an IBM 2012 word that refers to data unreliability and is usually linked with information based on people's emotions or any other type of information dependent on human judgment. According to the same authors, Oracle defined value in 2012 and linked it to the concept of essential data, indicating that high-value information can be recovered through extensive data analysis. According to Mathrani et al. [6], veracity and value constitute the rigor of big data analytics (BDA). They are significant because other qualities of big data processing, such

as volume, speed, and variety, would be useless without data analysis. Advanced analysis techniques are utilized to extract relevant information from large amounts of data, allowing for better decision-making. According to Gunasekaran et al., a variety of technologies, including sensors, barcodes, RFIDs, and Internet of things, are being utilized in supply chain management (SCM) to integrate and coordinate all the links in the chain (2016). The BDA is altering supply networks, and its use in supply chain management (SCM) has been reported on various issues. Since empirical evidence suggests that BDA in SCM can reduce operating costs and improve chain process agility, there is a growing interest in discovering the optimum sites and processes to implement BDA.

Several publications and literature reviews [7–9] and Mohammad Alsaffar et al. [10] examine big data analytic applications in the supply chain, with the vast majority focusing on operational settings such as factory production lines, product development, and product assembly, among other things. Chong [7]; Majidian et al. [8]; Solanki et al. [9]; and Abdulsattar Abdullah Hamad et al. [11] all reviewed the literature on material flow in industrial operations, with Borgi et al. [12] and Sengan and Set al [13] focusing on transport and logistics. According to Husamaldin et al. [14], few research studies have evaluated the supply chain from the perspective of BDA techniques. As a result, this study will employ a comprehensive mapping analysis and a survey to identify the supply chain regions where BDA has the most significant influence and develop a panoramic view of the current literature.

2. Theoretical Foundations

This section covers supply chain management, big data analytics, systematic mapping, and survey studies. These concepts are presented in such a way that they appear to be related to the project's purpose.

Supply chain management is the management of the supply chain. According to Awwad [15], the term supply chain is used more frequently in the literature than the definition of supply chain management. According to Leveling [16], supply chains are made up of firms that move commodities. Multiple organizations collaborate as part of a supply chain to form an end-to-end procedure for manufacturing and delivering a product. This chain includes raw material and component producers, product makers, wholesalers, retailers, and transportation companies. Varela Rozados et al. [17] concur on the definition of a supply chain as the alignment of firms that deliver products or services to market. Figure 1 displays various alternatives.

According to Felea et al. [18], SCM definitions fall into three categories: management philosophy, implementation, and management processes. Reference [19] suggests that supply chain management employs a systems approach, which views the chain as a coherent whole, rather than a collection of separate pieces. As a result, a new strategy known as supply chain management aims to coordinate the complete flow of commodities from maker to customer. It is a set of beliefs held by each organization in the supply chain that directly and indirectly impacts all other supply chain members and improves overall supply chain performance.

Albastroiu Nastase et al. [20] emphasize the importance of developing management practices that enable businesses to operate or behave in accordance with their management philosophy. As a result, several authors have written about supply chain management activities. They argue that businesses must widen their integrated behavior to encompass customers and suppliers to compete effectively today. Albastroiu Nastase defines supply chain management as the outer integration of integrated behaviors. In this context, supply chain management is defined as a set of actions that carry out the philosophy of chain management: chain management implementation. To meet the ever-changing expectations of the ultimate client, suppliers, operators, and manufacturers collaborate in a coordinated effort known as supply chain management.

In contrast to other authors, many authors, such as Lamba [21], emphasize the activities that comprise supply chain management. In other terms, a process is a collection of operations with a defined framework and measurable outcomes that collectively attempt to suit the objectives of a specific client or market. To put it another way, supply chain management, according to Saleem et al. [22] and Ahmed et al. [23], is a system for managing the movement of physical objects and associated data from the point of supply to the point of consumption to increase customer service and economic value.

Big Data Analytics. In contrast, Banik et al. [24] defined big data in 3 Vs in their definition of the concept. A new definition of big data was given by Lamba et al. [21]; it consists of five components: volume, speed, diversity, veracity, and value. Babiceano and Seker (2016) conclude from their analysis that the veracity and value aspects of big data, which represent elements related to the reliability and importance of information, are crucial for BDA because, without analysis, other big data processes would be of little value. Ramannavar [25] succinctly summarizes the BDA concept by referencing the use of sophisticated analysis techniques to extract useful information from vast amounts of data.

According to Husamaldin [14], huge data are useless unless that can be interpreted. When big data are leveraged to guide business choices, its true value can be seen. To support evidence-based decision-making, organizations need to be able to transform vast volumes of speedy and diverse data into valuable information. Extraction, cleaning, and annotation; integration, aggregation, and representation; modeling and analysis; and interpretation are the five basic steps in the process of obtaining information from vast volumes of data. These five steps are divided into two main subcategories in Solanki [9]: analyses of data management and information storage. Data management refers to the processes and tools used to gather, store, prepare, and retrieve data for analysis. In contrast, "analytics" refers to the techniques used to study and gain understanding from important data. Big data analysis can be seen as a smaller
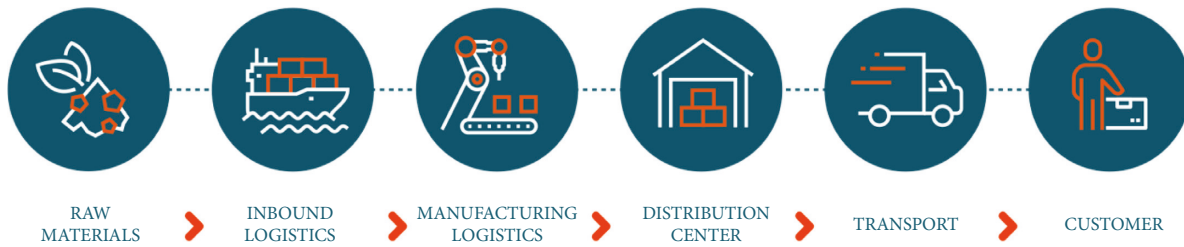
FIGURE 1: Supply chain.
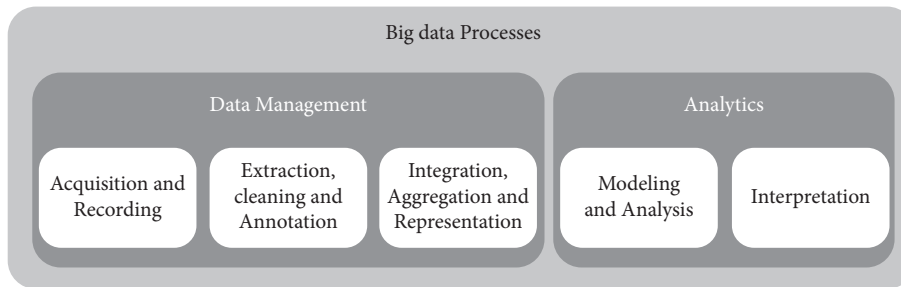
FIGURE 2: Big data processes [26].

phase in the larger process of extracting information from pertinent data, as seen in Figure 2.

Systematic Survey and Mapping Study (SSMS). According to Kapliński [26], SSMS aims to analyse a wider range of studies and rate the best representative studies on particular topics. In its study, SSMS primarily addresses issues connected to certain courses. Examples of these questions include those relating to research areas, empirical approaches, frequently utilized techniques, and the degree of automation of such techniques. On the other hand, surveys, according to Varela Rozados et al. [17], seek to understand the population from whom sampling was done. The company's largest developer population of 100 can be surveyed when 25 developers are questioned about a new procedure. Surveys are intended to produce generalizable findings. According to the authors, surveys may provide a huge number of elements to analyse. Still, this reduction in variables is necessary to estimate the greatest comprehension by the fewest variables because it makes the data collection and analysis process easier.

SSMS, according to Borgi et al. [12], offers a wider range of applications than other emergency medical services. The SSMS considers a large number of papers; however, only classification data are generated on these investigations. The classification and combining of research in SSMS are done using a preformatted classification technique. Researchers frequently elaborate on these categories, such as the type of technique, study location, publishing type, and research method employed, based on the information supplied in the papers. Habib et al. [19] contend that surveys can be used as a pre-study for a more in-depth investigation. Not providing an answer to the original research question can open up new opportunities for analysis. In contrast, surveys can be used to make statements about some populations, to explain how

well people understand certain subjects, and even as a pre-study.

For a number of reasons, SCM while operating an SSMS is pertinent to this topic. To start, the method ensures that the literature review is impartial, exhaustive, and verified. Second, it intends to locate, evaluate, and compile all pertinent studies by conducting a complete study of the literature on supply chain and BDA. It offers a summary of a research topic and indicates whether or not there are any subtopics with sufficient research to conduct other types of analysis as the third component. This study also includes classification and analysis of current literature on a thematic level, which may be used to identify research gaps and suggest some lines of inquiry for future studies. The results of this study will be used to guide future research toward addressing the gaps that have been found.

To get a more complete picture of whether the processes being kinetically evaluated are in line with the beliefs and attitudes of the general population, the data collected through the SSMS will also be compared to the survey results.

3. Methodology

Research Objectives and Questions. The goals of this work are twofold: (1) to organize the existing research on applying big data analytics to the supply chain into useful categories and (2) to identify the chain processes where BDA has been applied most frequently. Based on this, for the systematic mapping study, six research questions (RQs) were formulated; in this way, the various particularities of the objectives can be carefully explored. According to Ramannavar [25], RQs in mapping studies should be generic to discover

research trends over time and topics covered in the literature. Table 1 summarizes the PQs as follows.

In addition, a survey was carried out taking into account the main areas of application of big data analytics in the supply chain to identify, based on the experiences of professionals in the area, which of these have the greatest impact on the chain. Lamba et al. [21] cite that a survey is a system to collect information from or about people to describe, compare, or explain their knowledge, attitudes, and behavior. Based on this, the questions shown in Table 2 were developed.

3.1. Mapping Study. This section aims to present how the review of the mapping study will be carried out. The first subsection, search strategy, deals with the search strategy used to obtain relevant studies on the topic. The second subsection, inclusion and exclusion criteria, describes the ways used to filter the studies. The third subsection, filtering the works, deals with how the works were filtered throughout the process to obtain the most relevant studies.

Search Strategy. As a means to accomplish the desired outcome, we shall specify the number of search engines. To ensure that search engines provide a more comprehensive result set, we will develop techniques based on a string structure that defines both the keywords and some of their synonyms.

Before searching for articles, it is essential to identify an effective set of keywords to capture the synthesis of existing literature related to our research topic. For this, the keywords are as follows:

(1) Supply chain management/Supply Chain/Management/Manufacture/Sale/Inventory/Logistics/Requests/Transport

(2) Big Data/Data/Data Volume/Data Speed/Data Variety/Data storage

(3) Data Analysis/Data Mining/Descriptive Analysis/ Predictive Analysis/Prescriptive Analysis/Machine Learning

From the definition of the groups mentioned above, the following string was created (("Supply Chain" OR Manufacturing OR "Order Picking" OR Logistics) AND ("Big Data Analytics" OR "Predictive Analytics")). Inclusion and Exclusion Criteria. This section aims to establish exclusion and inclusion criteria used to filter studies retrieved from selected electronic databases. The following list specifies the exclusion criteria (EC) defined. EC excluded studies that

(i) EC1: the purpose of this section is to define the criteria that were used to pick relevant studies from specific electronic databases. A list of the defined exclusion criteria (EC) is as follows.

(ii) EC2: patent registration or early-stage projects where an overview and roadmap are presented are examples

(iii) EC3: no keywords from the search string were present in the title, and the title's meaning runs

counter to the inquiries posed in the aforementioned research questions

(iv) EC4: no part of the study topics was addressed in the abstract

(v) EC5: appeared in duplicate

(vi) EC6: did not address big data analytics or supply chain issues. On the other hand, inclusion criteria (IC) were used to add work to our sample.

(vii) IC1: articles, final papers, master's and doctoral theses, or even dissertations focused on big data analytics or supply chain

(viii) IC2: was published or disseminated

(ix) IC3: studies published or available in scientific journals, conferences, pages of research groups, or educational institutions

(x) IC4: published until December 2021

Method of Selection. This section will cover how we narrowed down the available research to identify the most relevant papers. According to Mathrani et al. [6], a larger number of articles may not be preferable if a smaller number of articles more truly represent the desired topic's population. The following filtering procedure is presented to demonstrate the strategies used to identify representative studies using this methodology:

(i) Step 1: begin your search. The search string is sent to digital libraries to collect search results.

(ii) Step 2: the first filter. Any disparities discovered in the search results are removed. The EC1 and EC2 will be used for this. Calls for conference papers, special issues of periodicals, patent specifications, and reports with no peer-reviewed material are all examples of inappropriately retrieved works.

(iii) Step 3: sort by title. Exclusion criteria, EC3, are used to filter studies.

(iv) Step 4: sort by abstract. Research is filtered with EC4 in mind. To eliminate papers whose content is unrelated to the key questions addressed in the research questions, Table 1 was used.

(v) Step 5: combine. All of the filtered studies from the previous phase are grouped in one location.

(vi) Step 6: remove duplicates. A study is often found in two or more libraries. EC5 is then used to remove all duplicates, guaranteeing that each study is unique.

(vii) Step 7: Filtering of research using EC6 in the entire text, omitting papers that are irrelevant to the topic at hand addressed in the research questions.

(viii) Stage 8: exemplary works. The final selection of the most relevant studies is defined.

Filtering. This section presents the execution of the filtering process defined in the previous section, where a total of eight steps were defined. Figure 3 illustrates the result obtained in

TABLE 1: SSMS research questions.

Research questions	Motivation	Variable
RQ1: In which areas of supply chain management is big data analytics being applied?	Discover which areas of supply chain management big data analytic approaches are being used extensively	Supply chain areas
RQ2: At what level of big data analytic analysis is used in these areas of chain management supplies?	List the main levels of *big data analytics* used for the supply chain management	Big *data levels*
RQ3: What types of big data analytic models are used in supply chain management?	Classify the *big data analytic models* used for the supply chain	*Big data models*
RQ4: What big data analytic techniques are used to develop these models?	List the main big *data analytic techniques used* in the models	*Big data techniques*
RQ5: What are the search methods used in the works?	List the most search methods used	Search methods
RQ6: Where were the studies published?	List the target vehicles used to disclose the results	Search location

TABLE 2: Survey questions.

Id	Question
Q1	Considering supply chain processes, do you consider important management from big data analytic techniques?
Q2	Considering supply chain processes, do you agree that big data analytics can assist in managing stock in transit?
Q3	Considering supply chain processes, do you agree that big data analytics can assist in vehicle routing (logistics)?
Q4	Considering supply chain processes, do you agree that big data analytics can assist in locating selection for installations?
Q5	Considering supply chain processes, do you agree that big data analytics can assist in selecting suppliers?
Q6	Considering supply chain processes, do you agree that big data analytics can assist in demand-driven storage?
Q7	Considering supply chain processes, do you agree that big data analytics can assist in real-time demand processes?
Q8	Do you agree that big data analytics can help reduce costs considering supply chain processes?
Q9	Considering supply chain processes, do you agree that big data analytics can help so that there is no lack of product in the gondola?
Q10	Considering supply chain processes, do you agree that big data analytics can assist in collecting orders? In the previous questions, relevant points for the supply chain were mentioned. Select 3 processes you think are the most important, that is, the processes in your perception
Q11	Have more impact along the chain

each step of the filtering process. The initial search for articles retrieved 5,437 articles, and then, the EC1 and EC2 were applied, with 2.55% (139 articles) discarded as impurities. Continuing the filtering process, the EC3 was applied to the other articles (5,298), with 91.77% (4,862 articles) being filtered through the title review. Then, EC4 was applied in a sample of 436 studies, where 63.30% (276 articles) were discarded after reviewing the abstracts. The remaining studies (160 articles) were gathered, and this combination generated a sample of 134 articles; that is, 16.25% of the articles were filtered at this stage. The next step was to eliminate duplicates, running EC5; this step reduced 28.35% (38 articles). Then, EC6 was applied, this stage of the process discarded 26.04% (25 articles). When examining 71 remaining works, some situations were observed in which the subjects were approached similarly: works produced based on previous articles. Thus, 29.57% (21 articles) were discarded. Finally, 50 articles were selected as the most representative of this study (Figure 3).

3.2. Survey. In addition to the mapping study, a survey was carried out, or in Portuguese, to identify, according to the participants' experiences and views, which areas and processes of the big data analytic supply chain have had the greatest impact. The following sections provide an overview of the preparation and execution of the survey.

Selection of Participants. The choice of participants was made through definitions of categories of specialists that make sense for the work, based on their practical and academic experiences, which is, therefore, the limiting factor for other professionals who are not inserted in this environment to participate in the study job. Participants must meet at least one of the following criteria:

(i) Be the owner/partner of a company operating in some area of the supply chain or of a technology company that provides products and/or services to companies operating in some area of the supply chain

(ii) Be the manager of a company operating in some area of the supply chain or of a technology company that provides products and/or services to companies operating in some area of the supply chain

(iii) Be a technology professional from a company operating in the supply chain area

(iv) Be a strategic planning professional and know the supply chain

	Initial search	Initial Filter	Filter by title	Filter by abstract	Combination	Duplicate Removal	Filter by Text	Work Represents
ACM Library	1039 (Filtered 0.38%)	1035 (Filtered 90.82%)	95 (Filtered 69.47%)	29				
Google Scholar	519 (Filtered 0%)	519 (Filtered 70.90%)	151 (Filtered 0.38%)	50				
IEEE Explore	135 (Filtered 0.0%)	135 (Filtered 86.66%)	18 (Filtered 0.38%)	9				
Inspec Database	1206 (Filtered 5.88%)	1135 (Filtered 99.82%)	2 (Filtered 0.38%)	1				
Science Direct	885 (Filtered 1.80%)	869 (Filtered 94.47%)	48 (Filtered 0.38%)	21	134 (28.35%)	96 (26.04%)	71 (29.57%)	50
Scopus	361 (Filtered 4.43%)	345 (Filtered 80%)	69 (Filtered 0.38%)	35				
Springer Link	1180 (Filtered 0.59%)	1173 (Filtered 96.16%)	45 (Filtered 0.38%)	9				
wiley Online	112 (Filtered 22.32%)	87 (Filtered 90.80%)	8 (Filtered 25%)	6				
Total (Filtered)	54.37 (2.55%)	5298 (91.78%)	436 (63.30%)	160 (16.25%)	137 (28.35%)	90 (26.04%)	71 (29.57%)	

Figure 3: Results of the selection process.

   (v) Be a business consultant and have experience with supply chain companies

   (vi) Be a professor or researcher on big data analytics/ supply chain.

Preparation of the Questionnaire. For the elaboration of the questionnaire applied to the study participants, the references acquired during the systematic mapping study were used, where the main points of impact for the supply chain were raised. From this, ten questions emerged to understand the relevance of big data analytics and also to compare the results between the opinions of professionals and the studies being developed in these areas. In addition to the questions directly related to the supply chain's areas of activity, another question was created for the participants to list, among the questioned areas, which, in their view, were the three most relevant when related to the use of big data analytics.

Planning and Execution. Given the construction of the questionnaire, it was necessary to find professionals who fit the study's participants' profiles. For this, it was decided to search for companies that act in this business segment and then make the questionnaire available to employees. A link was made available to answer the questionnaire through email and social networks to reach the participants.

4. Results

This section presents the results obtained through the mapping study and the survey in two subsections. In relation to the mapping study, each research question formulated in Table 1 will be discussed, and about the survey, the questions formulated in Table 2 will be addressed. For the results of the mapping study, the fifty most representative articles were taken into account. For the survey results, 25 responses obtained through the applied questionnaire were considered.

4.1. Systematic Mapping Study. RQ1: *Supply Chain Areas.* This question raises the areas of the supply chain where big data analytic concepts are used most frequently. The main results shown in Table 1 will be presented as follows. It can be seen that 78% or 39 of 50 articles evaluated focus on the areas of supply chain management and demand management. These results, concentrated in these two areas, are understandable, as they are areas where management and decision-making are constantly being exercised. BDA processes are applied to bring visibility and information to assist decision-making in these areas. In addition to this information, understanding the market's needs has been a very valuable differential; mastering this variable implies having effective control of production and avoiding large productions without demand, which consequently reduces large numbers of inventories and expenses with materials and unnecessary cousins. The opposite also applies in these cases, producing little for a lot of market demand. Representing 22% or 11 of 50 articles evaluated are the areas of manufacturing, transport/logistics, and storage/warehousing. These areas are directly linked to production and

not management; they are areas where, in general, the focus of the application of BDA is aimed at optimization, whether by reducing time, cost, raw material, and other variables.

It is visible that BDA is being applied to management and control areas, as having information in decision-making brings a great competitive advantage; however, despite being little explored, the production areas of the chain have a lot to gain from applications of BDA, as many processes can be optimized, adding value to the product.

RQ2: *Levels of Big Data Analytics*. This research question raises at which levels BDA concepts are being applied most frequently. The main results shown in Table 2 will be presented as follows. Among the levels that were found, 66% or 37 of 50 articles evaluated represent the levels of predictive and descriptive analysis. These types of analysis are related to each other, while predictive analysis seeks to find future movements based on the data being reported by the most diverse platforms; the descriptive analysis seeks to find relationships and/or associations between historical data, and from these data, it predicts future movements. The levels of mixed and prescriptive BDA represent 34% or 17 of 50 articles evaluated; these two levels are somewhat generic, as the mixed level deals with the level of BDA that uses more than one technique to obtain the results, while the prescriptive deals with of tools and/or mechanisms used for analysis and presentation of information obtained by BDA techniques.

It is possible to see the trend of BDA levels for predictive and descriptive analysis, the advancement of technology, and the way we connect with it, which has been advancing over the years, and along with it, the amount of information that is generated by the most various means has been increasing exponentially, enabling processes to be created to analyse this lake of information and interpret and predict situations.

RQ3: *Big Data Analytic Models*. This research question raises the most frequently applied BDA models. Together, the optimization and prediction models represent 48% or 24 of 50 articles evaluated. The optimization model, in its essence, seeks through information to optimize some stage of the chain, whether by reducing time, cost, or processes. In contrast, the forecast model seeks to provide forecasts for better positioning, which is generally related to management and decision-making. Classification and simulation models represent 32% or 16 of 50 articles evaluated. The classification model lists the chain's main processes to focus on the most relevant processes.

In contrast, simulation processes focus on simulating future situations, usually in manufacturing processes, as they can predict and/or avoid possible problems. The other models represent 20% or 10 of 50 articles evaluated; they are visibility, mixed, and others. Despite being the models with less prominence, they are models that relate to each other by presenting techniques related to information management, with a focus on presenting the data. They bring a fundamental role that has already been seen in other sections; however, they are still little explored in the form of a BDA model.

Analysing the most used BDA models in the articles, it is possible to perceive a greater representation of the optimization and prediction models. However, the main focus of the BDA has been the management of information and aid in decision-making, and these models can indeed be used for this purpose, but it would be interesting to analyse why visibility reasons are being little used.

RQ4: *Big Data Analytic Techniques*. This research question raises the most frequently applied BDA techniques. Mixed techniques, visualization, and heuristic approach represent 68% or 34 of 50 articles evaluated. These techniques have a high representation because the mixed technique, as the name implies, uses two or more techniques to buy a BDA structure. The visualization technique is usually used to complement data mining techniques, one of the most used techniques in recent years; we can find this technique in most articles with mixed classification. On the other hand, the heuristic approach is widely used for optimization, and as this is one of the most used BDA models, it is understandable to use this technique on a large scale. The other techniques presented represent 32% or 16 of 50 articles. These are techniques that, despite their importance, are linked to very specific processes and therefore appear in smaller quantities; they are usually used in conjunction with other techniques.

It is possible to perceive a tendency to use techniques that we can call generic, as they can be applied at all stages of the chain, in addition to being techniques focused on providing information and optimizing processes. At the same time, specific techniques have been little used, despite providing a great differential.

RQ5: *Research Models*. This research question seeks to understand the research methods used by the selected studies. According to Mathrani et al. [6], studies can be categorized into six research methods: solution proposal, evaluation research, validation research, opinion article, experience article, and philosophical article [15]. Evaluation research models and solution research models represent 66% of 50 articles; these research models suggest that articles are being developed to evaluate existing processes in the chain and/or propose new solutions for them. The other models represent 34% of 50 articles evaluated. They suggest opinion models and/or validations of concepts and techniques applied in the studies.

Finally, it is possible to notice that works related to the subjects studied, for the most part, aim to evaluate concepts and techniques that are being used or provide our solutions for the processes. This behavior tends to happen due to the new technologies available on the market, as they open up many possibilities for improvement in processes.

RQ6: *Research Location*. This research question investigates when and where the evaluated studies were published. The main objective is to examine the recent research trend of using big data analytics for supply chain processes. For this reason, the year of publication and the source for each selected study were collected. Figure 4 provides a chronology for all the studies evaluated. The solid gray line in Figure 4 summarizes the number of evaluated studies published per year from 2010 to 2021.
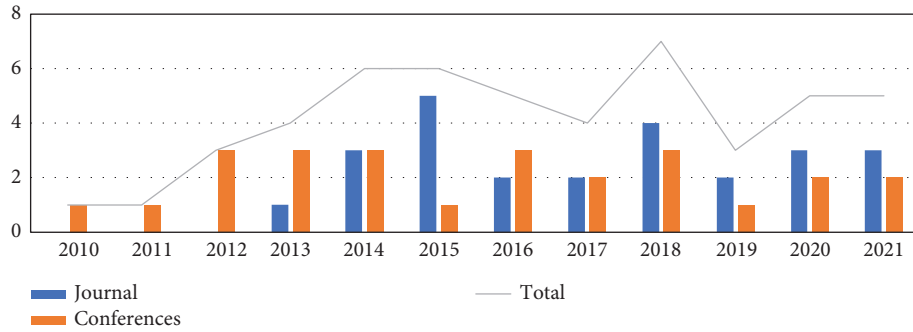
FIGURE 4: Research locations and year.

Although the period defined in the search strategy is until 2021, no studies were found before 2010. This may have occurred because big data had its first definition in 2001, but it was still very far from what we understand today; until one had a real understanding and usability of these emerging technologies, it may have taken a few years. Between 2011 and 2014, research on BDA for the supply chain experienced a considerable increase, but in 2016 with the great explosion of big data, research had the main peak.

It can be seen that in the years 2016 and 2021, the number of research related to the topic had considerable number. However, it will be necessary to follow the following years so that it can be affirmed that BDA has been widely used for supply chain processes since in 2021, where the polls suffered a slight drop.

4.2. Survey. Following the description of the data obtained, there is an analysis of the survey results so that it is possible to compare them with the results obtained in the mapping study. To focus on the analysis of the results, we will start with question number 11, where the results shown in Table 3 will be presented as follows. The main objective of this question is to list, among the areas of the supply chain, which participants consider BDA to have the greatest impact. Inventory management, stockout, and cost reduction together account for 67% or 50 of 75 responses, so questions regarding these three areas will be presented as follows.

Q1: *Inventory Management.* This question seeks to understand, in the view of experts, whether the supply chain inventory management process can be aided by the use of BDA. The main results shown in Table 4 will be presented as follows. Among the responses obtained, 84% or 21 of 25 responses fully agree that BDA can help this process, while 16% or 4 of 25 responses agree.

It can be seen that all participants agreed that this is one of the processes that can be most helped by the use of BDA; there were no neutral or negative responses. It remains a point of attention so that works and activities can be developed using this chain area.

Q9: *Disruption Management.* This question seeks to understand, in the view of experts, whether the supply chain disruption management and control process can be aided by the use of BDA. The main results shown in Table 5 will be presented as follows.

TABLE 3: Main supply chain processes.

Question 11	Answer number	Percentage
Inventory management	22	29
Lack of product in the gondola (rupture)	19	25
Cost reduction	9	12
Real-time demand processes	6	8
Demand-driven storage	5	7
In-transit inventory management	4	5
Vehicle routing (logistics)	4	5
Supplier selection	3	4
Order collection	2	3
Location for facilities	1	1
Total	75	

TABLE 4: Inventory management.

Question 1	Number of answers	Percentage
You fully agree	21	0.84
Agree	4	0.16
Neutral	0	0
Disagree	0	0
Totally disagrees	0	0
Total	25	

5. Discussion and Challenges

This study seeks not only to provide, through the mapping study, useful information to researchers and professionals interested in understanding which areas of the supply chain are most impacted by BDA but also to bring a perception of how these processes are being seen and used in practice, to create a relationship and address new challenges.

Perception of Academic Outcomes vs. Professionals. As can be seen through the interpretation of Tables 1 and 3, most academic results are focused on presenting BDA resources aimed at chain management, that is, methods that assist in decision-making in the most diverse processes; on the other hand, according to the sample carried out with professionals in the area, BDA has been used and/or in the perception of professionals, it can better help the areas of stock and rupture. This lack of connection between what is being researched and what is being applied in practice opens many doors for new investigations, whether focusing

TABLE 5: Breakage management.

Question 9	No. of answers	Percentage
Totally agree	21	84
Agree	3	12
Neutral	1	4
Disagree	0	0
Totally disagree	0	0
Total	**25**	

Bold value mean number of Total Answers.

research and efforts in areas where the concepts are already being used, or applying these concepts in management processes. Applying, executing, and monitoring some of the techniques in practice should bring interesting results to be analysed and disseminated.

Architecture for Real-Time BDA. In the course of this work, a strong trend in the use of predictive techniques and machine learning can be analysed, where the strongest characteristics are the use of historical data creating relationships; however, there is a deficiency in techniques for using BDA in time. This tends to be a prosperous path for the future, as concepts of smart cities and other advances in the most diverse devices make information increasingly complete and more accessible, making everything very dynamic. More and more information is obtained in less time, and it is necessary that analysis techniques can keep up with this pace.

6. Conclusions

The study's purpose was to categorize the available literature and then list the primary supply chain areas where big data analytic approaches are used. The notion of conducting a survey and generating a sample for comparison was developed to establish a relationship between what is being researched and examined in the academic community.

Using the filtering approach used for the mapping study, a total of 50 articles were selected for analysis. These publications were then categorized based on the research questions. These research questions enabled it to deduce the course of the investigation. More than 70% of BDA concepts were discovered to be employed in managing some process in the supply chain, with the vast majority of that time spent controlling that process and the remainder focusing on demand management. A survey of 25 professionals, however, reveals a very different picture. BDA is particularly valuable to professionals in these areas.

First, the current literature was studied and classified; second, the supply chain areas most affected by BDA were identified; and third, the strategies used to execute BDA were provided. To put it another way, it is commonly acknowledged that these contributions have a substantial impact because they direct study in these areas and bring a sample of the population's understanding to the table.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this study.

References

[1] D. Bumblauskas, H. Nold, P. Bumblauskas, and A. Igou, "Big data analytics: transforming data to action," *Business Process Management Journal*, vol. 23, no. 3, pp. 703–720, 2017.
[2] S. Yin and O. Kaynak, "Big data for modern industry: challenges and trends [point of view]," *Proceedings of the IEEE*, vol. 103, no. 2, pp. 143–146, 2015.
[3] P. Kaur, "Managing big data: a step towards huge data security," *International Journal of Wireless and Microwave Technologies*, vol. 6, no. 2, pp. 10–20, 2016.
[4] K. N. Aye and T. Thein, "A platform for big data analytics on distributed scale-out storage system," *International Journal of Big Data Intelligence*, 2015.
[5] N. Elgendy and A. Elragal, "Big data analytics: a literature review paper," *Advances in Data Mining. Applications and Theoretical Aspects*, vol. 8557, pp. 214–227, 2014.
[6] S. Mathrani and X. Lai, "Big data analytic framework for organizational leverage," *Applied Sciences*, vol. 11, no. 5, p. 2340, 2021.
[7] D. Chong and H. Shi, "Big data analytics: a literature review," *Journal of Management Analytics*, vol. 2, no. 3, pp. 175–201, 2015.
[8] S. Majidian, R. Vanani, and Iman, *Literature Review on Big Data Analytics Methods*, Intechopen, London, UK, 2019.
[9] M. S. Solanki and M. A. Sharma, "A literature review on big data analytics," *International Journal of Innovative Research in Computer Science & Technology*, pp. 234–247, 2021.
[10] M. Alsaffar, A. A. Hamad, A. Alshammari et al., "Network management system for IoT based on dynamic systems," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 9102095, 8 pages, 2021.
[11] A. A. Hamad, M. L. Thivagar, J. Alshudukhi et al., "Secure complex systems: a dynamic model in the synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9719413, 6 pages, 2021.
[12] T. Borgi, N. Zoghlami, and M. Abed, "Big Data for Transport and Logistics: A Review," in *Proceedings of the 2017 International Conference on Advanced Systems and Electric Technologies (IC_ASET)*, pp. 44–49, Chennai, Tamil Nadu, January 2017.
[13] S. Sengan, O. I. Khalaf, S. Priyadarsini, D. K. Sharma, K. Amarendra, and A. A. Hamad, "Smart healthcare security device on medical IoT using raspberry pi," *International Journal of Reliable and Quality E-Healthcare*, vol. 11, no. 3, pp. 1–11, 2022.
[14] L. Husamaldin and N. Saeed, "Big Data Analytics Correlation Taxonomy," *Information*, vol. 11, 2019.
[15] M. Awwad, P. Kulkarni, R. Bapna, and A. Marathe, "Big data analytics in supply chain: a literature review," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Washington DC, USA, September 2018.

[16] J. Leveling, M. Edelbrock, and B. Otto, "Big Data Analytics for Supply Chain Management," in *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2014)*, Selangor Darul Ehsan, Malaysia, December 2014.

[17] R. Varela, Ivan, Tjahjono, and Benny, "Big Data Analytics in Supply Chain Management: Trends and Related Research," in *Proceedings of the 6th International Conference on Operations and Supply Chain Management*, Wellington, New Zealand, 2014.

[18] M. Felea, A. Nastase, and Irina, "Defining the concept of supply chain management and its relevance to Romanian academics and practitioners," *Anfiteatro Economic*, vol. 15, pp. 74–88, 2013.

[19] Md. M. Habib, *Supply Chain Management (SCM): Theory and Evolution*, Intechopen, London, UK, 2011.

[20] A. Nastase, Irina, Felea, and Mihai, "Defining the concept of SCM and its relevance to Romanian academics and practitioners," *Amfiteatru Economic*, vol. 15, pp. 74–88, 2013.

[21] K. Lamba and S. P. Singh, "Big Data analytics in supply chain management: some conceptual frameworks," *International Journal of Automation and Logistics*, vol. 2, no. 4, p. 279, 2016.

[22] H. Saleem, Y. Li, Z. Ali, M. Ayyoub, Yu Wang, and A. Mehreen, "Big data use and its outcomes in supply chain context: the roles of information sharing and technological innovation," *Journal of Enterprise Information Management*, vol. 34, no. 4, pp. 1121–1143, 2020.

[23] A. S. Al-Obeidi, S. Fawzi Al-Azzawi, A. Abdullah Hamad, M. L. Thivagar, Z. Meraf, and S. Ahmad, "A novel of new 7D hyperchaotic system with self-excited attractors and its hybrid synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 3081345, 11 pages, 2021.

[24] A. Banik and S. Bandyopadhyay, "Big data-A review on analysing 3Vs," *Journal of Scientific and Engineering Research*, vol. 3, no. 1, pp. 1–4, 2016.

[25] M. Ramannavar and N. Sidnal, *Big Data and Analytics—A Journey through Basic Concepts to Research Issues*, Springer, New Delhi, India, 2016.

[26] O. Kapliński, N. Košeleva, and G. Ropaitė, "Big Data in civil engineering: a state-of-the-art survey," *Engineering Structures and Technologies*, vol. 8, pp. 165–175, 2016.

Research Article

Blockchain Framework for Secure COVID-19 Pandemic Data Handling and Protection

Arshad Ahmad Dar [ID],[1] **Malik Zaib Alam** [ID],[1] **Adeel Ahmad** [ID],[1] **Faheem Ahmad Reegu** [ID],[1] **and Saima Ahmed Rahin** [ID][2]

[1]*Department of Computer Science and Information Technology, Jazan University, Jazan 45142, Saudi Arabia*
[2]*United International University, Dhaka, Bangladesh*

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

COVID-19 pandemic caused global epidemic infections, which is one of the most severe infections in human medical history. In the absence of proper medications and vaccines, handling the pandemic has been challenging for governments and major health facilities. Additionally, tracing COVID-19 cases and handling data generated from the pandemic are also extremely challenging. Data privacy access and collection are also a challenge when handling COVID-19 data. Blockchain technology provides various features such as decentralization, anonymity, cryptographic security, smart contracts, and a distributed framework that allows users and entities to handle COVID-19 data better. Since the outbreak has made the moral crisis in the clinical and administrative centers worse than any other that has resulted in the decline in the supply of the exact information, however, it is vital to provide fast and accurate insight into the situation. As a result of all these concerns, this study emphasizes the need for COVID-19 data processing to acquire aspects such as data security, data integrity, real-time data handling, and data management to provide patients with all benefits from which they had been denied owing to misinformation. Hence, the management of COVID-19 data through the use of the blockchain framework is crucial. Therefore, this paper illustrates how blockchain technology can be implemented in the COVID-19 data handling process. The paper also proposes a framework with three main layers: data collection layer; data access and privacy layer; and data storage layer.

1. Introduction

1.1. The COVID-19 Pandemic. There are a plethora of coronaviruses that can infect people, along with the "novel" one that causes COVID-19. Animals are likely to be here for a considerable time. A virus involving animals can frequently infect humans. Experts claim it is what occurred inside. In other respects, although this virus is not unique to the globe, it is surprising to people. In 2019, when it began to horrify humans, researchers analyzed it as a novel coronavirus. Experts refer to these genotypes as SARS-CoV-2 was the first type known, also called as Alpha-Variant-B.1.1.7 detected in the UK, following the other known type was a Beta variant that was reported first in South Africa, the third variant was gamma reported in Japan, followed by the Delta variant reported in India in late 2021. The currently existing variants are Omicron-B.1.152a caught in South Africa with the new variants known available as on July 2022 include BA.4 and BA.5 variants of COVID-19. Hence, as a consequence of the coronavirus (COVID-19) epidemic that occurred in late 2019, there was an international health crisis. There were more than a million new cases throughout the globe in little over three months, and the virus had spread to various nations. Some of the most affected cities, like Wuhan, New York, Texas, and Hawaii, were reporting new cases by the hour during this period. The number of fatalities also continued to rise due to measures like lockdowns and social isolation, enacted as a response [1].

The cancellation of major international events like the Tokyo Olympics and the Dubai Expo due to COVID-19 has

also had a severe effect on the global economy. Therefore, from the first few cases reported by the Globe Health Organization, governments, research centers, and institutions all over the world have been working feverishly to produce a vaccine and forecast the predicted development of the coronavirus.

The daily number of positive and negative cases, tests done, individuals admitted to the hospital, casualties, recorded, space available in ICUs (both empty and occupied), ventilator shortages, and the number of medical experts on-site all play a role in effective containment strategies. COVID-19's development may be monitored and judgments made based on these statistics. As a result, the pandemic of COVID-19 is also known as a data-driven one.

Data availability, verification, and integrity in the COVID-19 pandemic management are crucial for making precise and effective conclusions and recommendations to the public based on records and statistics. For this, tracking applications developed by governments have been adopted to stop the spread of the virus and to ensure the integrity of data. A patient's development following clinical trials may be tracked via tracing applications, making them a key tool for vaccination performance monitoring. Moreover, health facilities use traditional tools such as Excel and centralized databases to store and analyze the data [2].

Governments and significant healthcare facilities have had difficulty managing the pandemic in the absence of appropriate drugs and vaccines. It is also very difficult to track down COVID-19 patients and manage the data produced by the epidemic. Managing COVID-19 data presents additional challenges in terms of data privacy, access, and collecting. With the use of blockchain technology, people and organizations will better manage COVID-19 data thanks to features like decentralization, anonymity, cryptographic security, smart contracts, and a distributed architecture, that is, the miraculous gift of blockchain technology providing the novelty in pandemic data handling and confidential aspects.

1.2. Opportunities to Current COVID-19 Data Handling. Blockchain is still a relatively new technology for managing data electronically, but evidently, it has great potential for accountability and supports transparency substantially. A blockchain manages a ledger of transactions, making identical copies for the members to view over the computer network altogether. It implies that all transactions done by the users of the blockchain framework are recorded in some type of database termed as the ledger. Due to its decentralized nature, the similar copies of this information recorded in the above discussed database enable every user to keep the accountability of the electronic data stored in these ledgers. Considerably, blockchain is ideal for transactions that are lightweight in terms of their digital footprint, with the advantages of immutability and greater transparency as well. During COVID-19, the healthcare systems were overwhelmed to a major extent. Blockchain technology plays a crucial role in identity verification, managing the medical supply chain, and with the patient's

consent, sharing and accessing patient-specific data throughout the channel [1].

Blockchain-enabled tools have emerged to fight the global pandemic in the health sector like an identity management system to support contact tracing in the South Korean region. This system also made the sharing of Covid-specific data available for research purposes as well. Previously, blockchain has been used for supply chain management of pharmaceuticals, for the supply of medicines, adequate medical supplies, and for coming up with appropriate vaccines to combat the deadly virus on the whole [3]. In Figure 1, the potential uses of blockchain technology specific to the healthcare sector are highlighted:

As it turns out, Blockchain is not an ideal option to store high-volume data due to its capacity limitations and on the computational front, it is difficult to replicate data across every network. Moreover, research has found out that storing large records, like electronic medical records or records relating to genetic data, is quite costly over a blockchain and it is inefficient on the whole. It is also very difficult to post query data on a blockchain, which eventually makes the data unavailable for any purpose at all.

Global efforts were made during the pandemic to develop and test vaccinations to stop the viruses' transmission and identify carriers as quickly as possible to find the most effective remedies. Blockchain has the potential to fully satisfy various requirements like security and protection of data, ease of data sharing, and easy access to a vast amount of data as well. During COVID-19, the field of medicine tried taking maximum advantage of blockchain technology through clinical trials and other significant processes involved in the treatment of the virus. The main use of blockchain technology was to manage the epidemic in an effective manner, not only just in the healthcare sector but also in providing adequate solutions for other purposes during lockdowns. Blockchain was evidently very crucial during the COVID-19 times as it offered reliable, transparent, and relatively low-cost solutions to corporations for effective decision-making. The quickness offered through blockchain technology helped many during this crisis and with the passage of time, it is becoming a significant and more integral part of the fight against Coronavirus [4].

The main motive behind the intervention of blockchain technology was to enable easy tracking and effective monitoring, ensuring an adequate supply chain of products and services during COVID-19, and securing the online payments to a major extent as well. All this is possible because blockchain technology is able to create a chronological order of encrypted signatures, a secure ledger that is easily shared by all the members present over the network. However, due to the limitations posed by the infrastructure, on the whole, blockchain technology is not able to fully take advantage of this state-of-the-art unique technology [5].

In comparison with traditional database systems, blockchain technology makes use of its inherent quality to ensure data gathering and management, transactions are transparent, immutable, and accurate due to its decentralized approach, data accessing and storing is much more rapid and does not need the control of central authority as
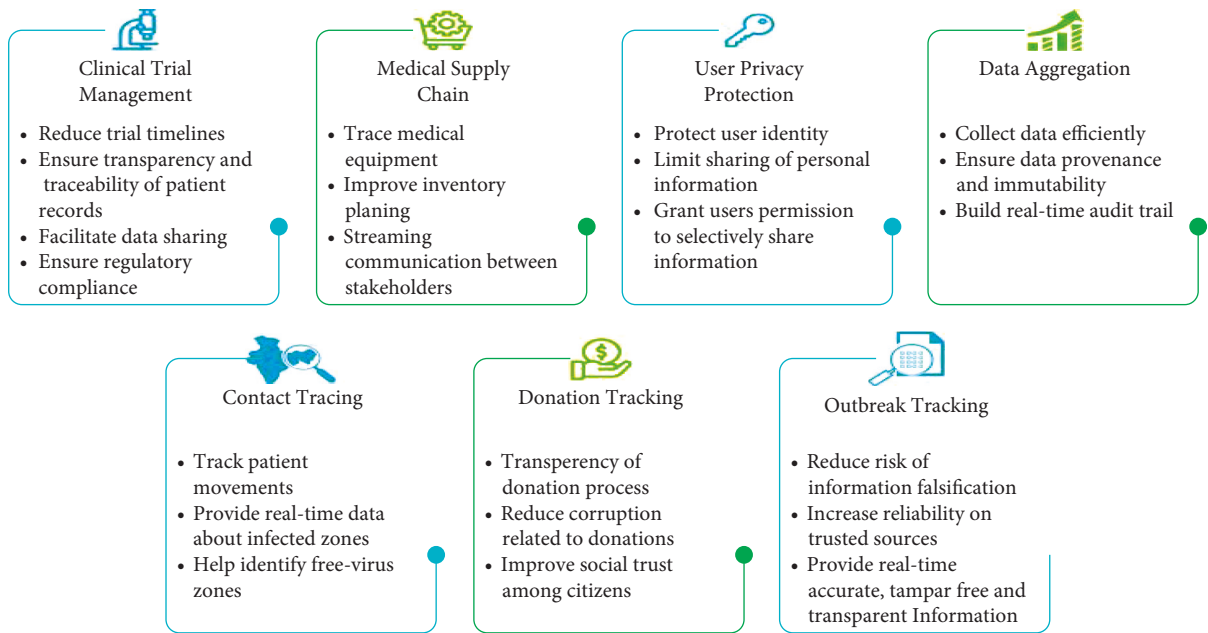
Figure 1: Use of blockchain technology.

done in the centralized approaches like the DBMS framework where the management transactions are not transparent and require the central control. Blockchain also makes it possible for multiple parties to easily connect in a digitally enabled environment and exchange money without the need for a central authority. Blockchain is transforming a wide range of industries in a variety of ways by easing value exchange and transparency, and boosting confidence across corporate systems. It is used in numerous fields, including finance, legislation, supply chain, travel, and healthcare sector, as it promises to improve data privacy in the healthcare sector and ensure safe data management. As a consequence, it is ideal for dealing with coronavirus-related healthcare issues as well [6].

The COVID-19 outbreak is a knowledge crisis because the available facts are vital in being preventive and protective. However, inaccurate statements that yielded untrustworthy and misleading data greatly expanded the agony of the pandemic. The actual prevalence of the ailment, the proportion of fatalities, indicators, potential therapies, and some of the most effective pandemic control measures have all been distorted by media platforms. The epidemic has aggravated the crisis of conscience in the administrative and medical centers as a result of the collapse to supply of timely and accurate facts about the current predicament. Hence, it leads to the requirements of Covid- 19 data handling to provide all benefits to the user in which they suffered due to misguidance, and therefore to acquire certain features like data security, data integrity, real-time data handling, data management, etc. COVID-19 data handling by the employment of blockchain framework plays a significant role.

This paper highlights the changes that could be made to privacy laws like the Health-Insurance-Portability and Accountability-Act HIPAA to facilitate the seamless and friction-free exchange of patient health information between healthcare organizations that must work together to treat patients as well as the exchange of that information with researchers looking into ways to lessen its effects [7]. The primary aim of the plan is to evaluate the lung computed tomography outcomes of pediatric MPP and MP coupled with streptococcal pneumonia. Clinically, the combination illness of MP and SP is very frequent, and young respiratory doctors must conduct crucial studies on how to diagnose this kind of mixed pneumonia [8]. This paper's method evaluates these photos and summarizes them as positively or negatively for COVID-19 employing existing deep learning models such as VGG19 and U-Net, which have been inspired by new findings that link the occurrence of COVID-19 to findings in Chest X-ray images [9]. In this paper, we construct a deep learning approach to detect COVID-19 using chest X-ray scans and features extracted from the images. ResNet50, InceptionV3, and VGG16 are three prominent models that have been adjusted on an enhanced sample that was constructed by collecting COVID-19 and conventional chest X-ray scans from various database searches [10].

Therefore, from the recent trends available in the study for the detection of COVID-19, we can conclude that before the intervention of blockchain, computed tomography scans, X-ray scans, CNN methods, deep learning methods like ResNet50, VGG19, InceptionV3, VGG16, U-Net, etc., were employed for critical COVID-19 data collection and management.

1.3. Challenges with Current COVID-19 Data Handling. However, although the above methods have proven to be effective, various key challenges still limit governments and health institutions from effective containment of patient tracing and developing vaccines for coronavirus.

1.3.1. Data Security Issues. The notion of security encompasses limiting malicious entities from introducing false negatives and false positives into the system and ensuring data integrity. In the context of COVID-19 patient tracing, malicious entities may opt to inject erroneous entries into the data set or cause a denial-of-service attack. These forms of attack are more common with centralized servers, which currently are being implemented by a majority of health facilities.

In most cases, a centralized server is considered secured and trusted. They are responsible for storing a patient's personal identification information (PII) and managing the security keys used to decrypt and encrypt data. However, this poses a great risk to patients' data in case the server is breached. The centralized server acts as a single point of failure and access.

Moreover, centralized servers are concerned with keeping third parties *f* database. In the case that a malicious party is an authorized party on the server, this poses a greater risk to the patient's data since it can easily be leaked, deleted, or altered, resulting in erroneous analysis, statics, and ineffective recommendations to the public [11].

1.3.2. Trust and Data Integrity Issues. The COVID-19 pandemic is an information crisis as the mitigation and preventive measures highly depend on the data available. However, the suffering of COVID-19 has been greatly exacerbated by misinformation leading to unreliable data and inaccurate data. Social media has led to a lot of confusion about the actual prevalence of COVID-19, the number of deaths, symptoms, expected treatments, and some of the best strategies to control the pandemic [12].

Due to the failure to provide timely and accurate data about the current state, the pandemic has worsened the crisis of trust in government and health institutions [13]. Many individuals are looking into other ways to enhance trust within the healthcare system.

1.3.3. Lack of Real-Time Data Recording and Sharing Systems. Global health data synchronization is a key factor to combat the COVID-19 pandemic. Sharing important data such as the number of cases, the positivity rate, the health facilities available, and the progress in clinical trials must occur in real-time to keep the public aware of the current state of the pandemic and to support immediate response to the pandemic.

As mentioned, two particular data sets collected in Wuhan led to different implications due to a mismatch in the number of cases recorded. The first data recorded 425 early cases, while a different data set in Italy and Wuhan recorded a larger number [14].

1.3.4. Isolated and Disintegrated Data Points. To coordinate effective response measures, it is crucial to integrate siloed disparate COVID-19 data types from different nations and health facilities. However, this has been a long-standing challenge in the healthcare industry due to incompatible communication standards in EHR and different jurisdictions in the public sector. The private sector on the other hand brings the challenges of competition and proprietary systems. Therefore, most private health facilities choose to keep the data to themselves to maintain security of their data. While most hospital websites use well-designed websites to display their COVID-19 cases, higher health authorities such as WHO have found it hard to integrate this data or prove its integrity of the data [15].

1.3.5. Data Mismanagement by Intermediary Institutions. Currently, most countries have a pandemic reporting system in which hospitals and clinics diagnose and report to higher national authorities, which in turn report to a higher authority such as the World Health Organization. In cases where there are numerous intermediaries that the records have to pass through the reporting time may result in the loss of some of the data as the records are transferred from one institution to another due to data mismanagement. Furthermore, the use of traditional communication protocols between institutions exposes the data to hacking and security attacks, thus making it more difficult to detect altered data after a hack [16].

1.3.6. Lack of Transparency. COVID-19, like all other pandemics, relies heavily on monetary and in-kind donations to help health facilities and research institutions fight the pandemic. However, throughout the COVID-19 pandemic in 2020, numerous corruption cases have been reported [17]. This can largely be attributed to the centralization and "blackbox" aspect of the current logistics, warehousing, expenditure, and distribution of funds in health facilities.

Therefore, innovative technologies like Artificial Intelligence, Internet of things, and blockchain can be used to alleviate some of these issues. Blockchain in particular provides key features that would revolutionize how governments and health institutions handle COVID-19 data. However, there are security concerns with respect to blockchain technology such as data vulnerabilities and double-spending activities present in blockchain consensus protocols. And yet, there are still a variety of challenges like legal issues, latency issues, resource utilization issues, and widespread implementation issues that have to be resolved before the blockchain can be fully implanted to counter COVID-19 and enhance COVID-19 preventive measures.

1.4. Contribution. Blockchain technology is a decentralized technology with an immutable information structure, transparency, and enhanced security through cryptographic encryption. It is a decentralized ledger technology, that is, distributed and stores data in blocks chained together by hash functions [18].

1.4.1. Features of Blockchain Technology

(1) Decentralization. The technology offers decentralization by dis-intermediation of central entities and offers equal opportunity in decision-making to all peer nodes in the network.

(2) Consensus Mechanism. It is a fault-tolerant decision-making mechanism that allows peer nodes to reach a conclusion fairly on a network.

(3) Distributed Ledger. Any transaction data recorded on the blockchain is distributed to all authorized peer nodes. Each node can view and verify the authenticity of the data and its source.

(4) Immutable and Tamperproof Ledger. Once a transaction is committed to the blockchain, it is almost impossible to alter the data. Moreover, any alterations to the data will have to be agreed upon by at least 50% of the nodes in the network. This leads to tamper-proofing of transactions recorded on the blockchain.

(5) Smart Contracts. A smart contract is a self-executing piece of computer code that encodes the terms and conditions of an agreement in a legal agreement within the code. Any node on the blockchain network can initiate the smart contract if certain preset conditions are met [19].

This paper looks at how the key features of blockchain technology can be implemented in the healthcare industry to allow for secure and efficient data handling of COVID-19 data. Table 1 shows the comparison between the traditional client-centered centralized platform and blockchain platforms.

1.4.2. Layout of the Present Study. The various modules of this paper are organized as follows: Section 1 contains an introduction followed by Section 2 which details some of the use cases of blockchain technology for COVID-19, Section 3 proposes a framework that can be implemented to allow secure data handling with regards to the COVID-19 pandemic, Section 4 details some of the expected challenges in implementing blockchain technology, and Section 5 presents the conclusions of the paper.

2. Use Cases of COVID-19 Based on Blockchain

Improved data sharing, patient tracking, and clinical trials might all benefit from the adoption of blockchain technology in the healthcare sectors impacted by the coronavirus epidemic. An in-depth analysis of the existing research on the potential use of blockchain technology in the battle against COVID-19 may be found in this section.

2.1. Contact Tracing. COVID-19 has an average incubation period from infection to symptoms of about 5.5 days, and it is also estimated that most cases are symptomatic. Therefore, keeping track of the current positive cases and who they get in contact with is very crucial. Governments and health care facilities have been using mobile apps to trace patients, but there are still some existing challenges as mentioned above.

By enhancing the quality and dependability of data collected via mobile apps, blockchain technology adds an extra degree of security. Since blockchain has a write and read model, the data collected from patients is immutable.

Blockchain technology also monitors patients' conditions, virus-prone areas, and safe areas and updates this information to the distributed ledger in real-time. This ensures that government and health institutions are aware of the current state of an area at a particular time.

To offer these services, blockchain technology also incorporates other technologies such as AI, IoMT, and geographic information system. It means blockchain technology employs read-write model to keep track of data gathering and patient monitoring, and it can also assist us in obtaining the information related to areas that prevail under the safe zone or virus zone and hence keep the distributed ledger to update all these details. Moreover, it not employs blockchain technology as the fundamental base but also unites other methodologies like AI, IoMT, and GIS for enabling us with these features. There, blockchain technology provides a practical approach to tracing COVID-19 cases and provides a reference point for making future containment decisions and recommendations [20].

2.2. COVID-19 Database Integration. Data collection, accumulation, access, and information exchange among health institutions is a major opportunity for a worldwide response to the epidemic. Blockchain technology provides the opportunity to ascertain data integrity through its immutable features. The architecture of blockchain acts as a single ledger that can be implemented in the healthcare sector to record COVID-19 data. Any data collected from testing and clinical trial data is uploaded and stored on the blockchain. Authorized researchers and institutions can access the data and utilize it to further their research towards developing a cure.

Blockchain also provides four key frameworks, public blockchain, private blockchain, consortium blockchain, and hybrid blockchain [21]. In the above discussed frameworks, the first one is entirely distributed, with no requirement of permission, and available to everyone who joins. Blockchain systems enable all nodes to still have identical access, the opportunity to add new data items, and the potential to certify existing chunks of data. The second one is the restricted blockchain systems, also known as administered blockchains, which are governed by a singular entity. Who is permitted to be a node on a permissioned blockchain is regulated by the central authority. Additionally, the central body may not always accord every node an equal right to execute certain duties. Instead of being regulated by a specific organization, as with a private blockchain consortium, blockchains are permissioned blockchains overseen by the group. As a result, this framework of blockchains is more global than private blockchains, which boosts their resilience. Blockchain systems that are managed by a singular body but have some supervision offered by public blockchains, which would be necessary to carry out specific transaction approvals, are known as hybrid blockchains. Besides public blockchain, the later frameworks empower compliance management and audibility by providing digital keys to grand flexible and secure data sharing through different managerial levels.

TABLE 1: Comparison between traditional client-centered centralized platform and blockchain platform.

Aspect	Traditional platforms	Blockchain platform
Data handling	Supports read, write, amend, delete and update operations	Only allows read and write operations
Integrity of data	It's possible to tamper with or change data.	Data is immutable and easily auditable
Privacy of data	Highly susceptible to cyber-attacks and PII leaks	Data is stored using cryptographic hash functions that make PII anonymous
Transparency of data	Databases are not transparent or easily auditable	Data is available and stored in distributed ledger
Data security	Single-point failures or assaults provide a high risk.	Distributed ledger hence highly fault-tolerant

2.3. Clinical Trial Data Management. Clinical trials are very crucial when developing a cure to the coronavirus pandemic. However, every product should be thoroughly tested to ensure that it matches the required standards and note the possible side effects of the products. Clinical trials are carried out in three phases, of which the third phase incorporates the largest number of test participants, which makes it a challenging and resource-intensive phase.

Blockchain technology helps ensure the safety and privacy of the data. Additionally, since it has a distributed framework, the technology distributes the resources required to manage the network to various health institutions. As a result, this encourages trial data sharing and ensures regulatory compliance within the healthcare facilities [22]. It also creates a continuous and transparent medical record that allows researchers access to the trial data at any point in the trial process.

2.4. User Privacy Protection. Health care facilities and government agencies are finding it more difficult to strike a balance between the need to gather enormous amounts of data and the need to protect patient privacy. Data may be collected and stored in a more secure manner using blockchain technology. Within a blockchain network, patients are provided with unique identification numbers known as addresses. Blockchain addresses are anonymous and cannot be linked back to the patient's personal information or identification. However, any information transaction initiated by a patient can be easily verified on the blockchain using Merkle trees [23].

The user's privacy protection in the blockchain network is acquired with the assistance of secret keys. The usage of encryption keys is really a crucial component of privacy protection in blockchain technology. Asymmetric encryption is being used by blockchain networks to authenticate client interactions. Every client on these platforms has a pair of keys, essentially a public/private key, respectively. These keys are interconnected cryptographically and comprised of random data. A user can never logically extract another participant's secret key from their public key pair. Clients are guarded against hackers and protection is improved dramatically [24].

As a result, COVID-19 patients can easily upload their records to the blockchain or even share them publicly without the fear of exposing their identity.

2.5. Secure Data Storage and Sharing. One of the key advantages of blockchain technology is providing verifiable and secure data by using its distributed and decentralized features. During the COVID-19 pandemic, blockchain is instrumental in recording and storing the patient's data such as test results, symptoms, location, and medical history of the patient in a secure and immutable manner.

3. Proposed Blockchain Framework

In this section, the paper proposes a framework with four layers in the system and the interactions within the data handling process within the system. Figure 2, labeled as "proposed blockchain framework that can be implemented with COVID-19 data handling" depicts the main three layers under the headings of data collection: which aggregates the data from smartphones, hospital devices, and other data assembling equipment's, data access and privacy layer: it is a supporting element of the framework since it regulates whom or how can we access the data. Any data included in smart contracts is timestamped, and the location is likewise logged and data storage analysis: the patient data in this architecture is kept in decentralized databases that are contributed by various hospital nodes and user devices. Active medical equipment serves as a peer-to-peer node and provides data storage. All these modules are further detailed in-depth under Sections 3.1, 3.2, and 3.3, respectively.

3.1. Data Collection Layer

3.1.1. User Interface Layer. At the user interface layer, there are three components, smartphones, hospital devices, and GIS system. Within this layer, the system collects the patients' COVID-19 status and their geographic location. The patient's data were recorded in two states, user positive or user negative. The geographic location of the patient is the health facility or clinic at which the patient was tested. The framework requires two data entry points to maintain the honesty of a patient's record. The first data entry point is from the patient's smartphone while the second entry is from the health facility. The two data points have to be consistent; otherwise, they will be rejected by the smart contract layer [16].

3.1.2. Mobile Service Layer. The mobile service layer is a core component of the framework as it hosts decentralized applications (DAPPS). DAPPS layer contains decentralized
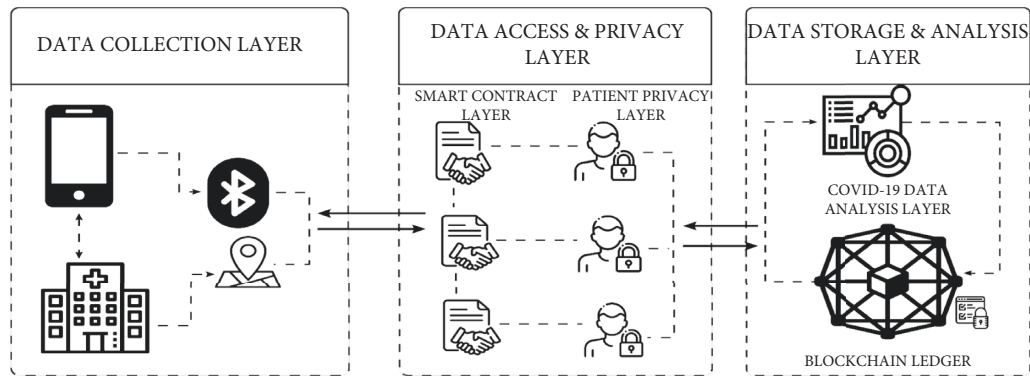
FIGURE 2: Proposed blockchain framework that can be implemented on COVID-19 data handling.

applications that collect data from the patients and interacts with the smart contract layer. The mobile service layer provides users and health facilities with crucial services: contact tracing based on Bluetooth, GIS systems, and health tracking services supported by data from the blockchain data storage layer [1]. In Figure 3, you can see the DAPPS Layer that enables users to access information on their smart devices.

3.2. Data Access and Privacy Layer: Smart Contracts.

The smart contract layer is the secondary component of the framework as it governs how and who accesses the data. Any record integrated into the smart contracts was time-stamped and the geographic location is also recorded from the highest level, country level to the smallest level, state level (depending on the country). Each contract will inherit data from a lower level contract, but two similar data points from a single patient will not belong to a single superior contract [25].

Moreover, data access will be restricted based on permissioned blockchain. The framework will provide various networks for users and the hospital to utilize. Within the network, numerous channels will be created to allow secure data sharing. The smart contracts will also implement unique identifiers for the patients to maintain their identity. Any patient's data stored will be recorded as follows: location status, infection status, and think progress. The contract will also be active for 14 days after the patient tests positive. Only after 14 days will patient data be considered clean status [26].

To save the operation cost of data storage, the patient data are hashed and only the head of the block is stored within user devices, the rest of the data is stored on the data storage layer.

3.3. Data Storage and Analysis Layer.

Within the proposed framework, the patient's data is stored within decentralized databases provided by different hospital nodes and user devices. Active hospital devices act as peer-to-peer nodes that provide storage for the data. The active nodes are incentivized to participate in the network by being allowed access to blockhead data. Since Bitcoin and Ethereum blockchain database designs have high computational costs

and slow transaction speed. Moreover, it is a peer-to-peer (P2P) network that can be used to transmit Bitcoin, a virtual currency, without the necessity for centralized power. In 2008, somebody or a body of citizens named Satoshi Nakamoto designed it. An irreversible decentralized system contains all transactions. On the other hand, blockchain-based shared platform Ethereum. Where ether is the name of Ethereum's network currency, the transactions in this instance are also kept in an immutable decentralized ledger. Therefore, the Hyperledger blockchain can be implemented in the framework to provide a hybrid blockchain platform as shown in Figure 3.

3.4. Effective Uses of Blockchain Technology during COVID-19.

When COVID-19 had hit the world in the year 2019 and took a stroll all over the world in the following year, the country struggled to gain control over its spread, be it the first world nations or third world nations. The countries were on the lookout for innovative and quick technologies that could help in managing the large amounts of data accumulated with respect to the Covid virus. The healthcare sector was one of the primary sectors to have been affected by this pandemic the most as it had no experience in dealing with such viruses ever before. Where at one end the healthcare sector was dealing with the large influx of patients, it had to be clinical trials to get to the root cause of the virus and come up with an effective remedy for it. In order for the clinical trials to run smoothly, it required a management system that was transparent and fair by all means. The countries were clear about the priorities of this system it had to be cost-effective since already the healthcare sector was facing tremendous expenses due to this global pandemic. Moreover, it was required to be fast, transparent, auditable, and compliant with the regulations everywhere. The following section will give a detailed overview of the application of blockchain technology in the form of effective solutions within the healthcare sector to gain control over the spread of COVID-19 [27].

3.4.1. Clinical Trial Management.

The main advantage of incorporating blockchain technology into this system was that it enables medical practitioners and researchers to gain
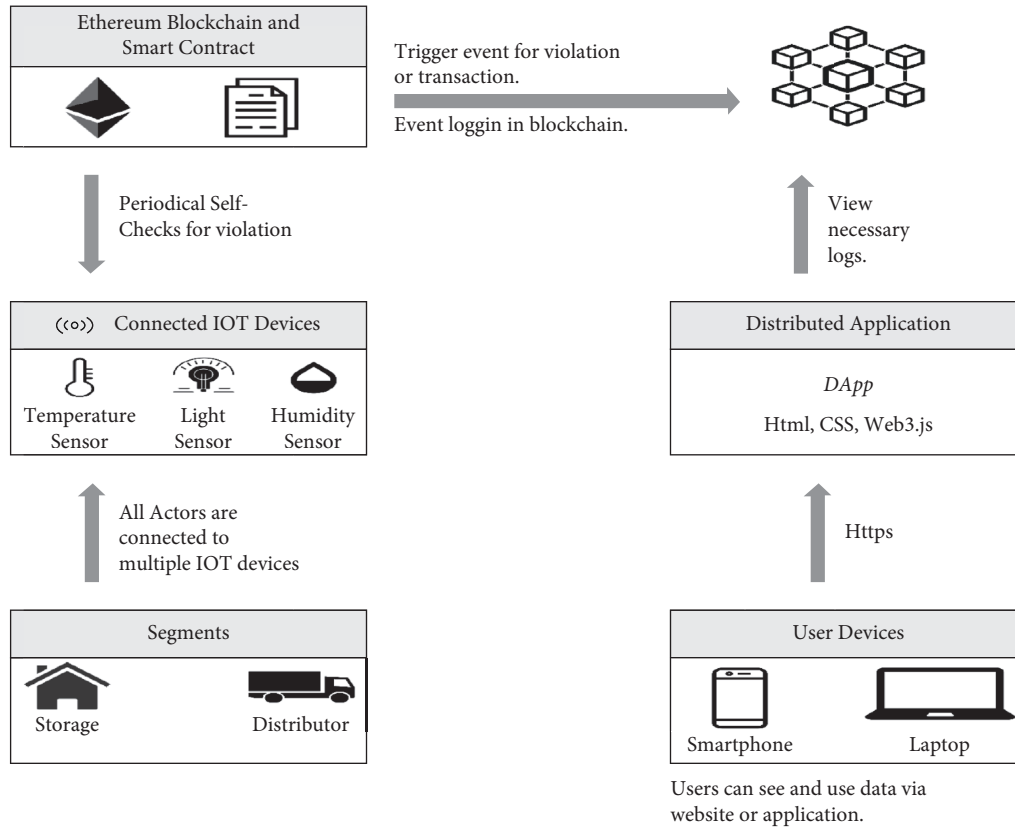
FIGURE 3: DAPPS Layer that enables the users to access information on their smart devices.

access to clinical data in a timely manner with greater accuracy altogether. One such application was introduced by the Canadians where it used blockchain technology to assist the local authorities and help the government bodies to a major extent in gaining control over the pandemic. Moreover, this application is also facilitated in clinical trials in relation to COVID-19, it associates the data with the individual person's ID, but the corresponding blockchain also records anonymous data inputs ensuring that the patient's sensitive information is not easily available to all. This technology can help in keeping track of the infected patients based on the data in the application this was a significant step toward ensuring that the spread of the virus was minimized to a major extent. It can further enable doctors to track the overall progress of the patients if they are infected by a coronavirus. The doctors can interact and share medication procedures to help out patients, who have isolated themselves at home. However, this application has some sensitive points that can question the sanctity and genuineness of the data shared on this blockchain technology [28].

3.4.2. User Privacy Protection.
During COVID-19, where all sectors were posed with major challenges left, right, and center, there was a dire need to strike a balance and that was possible through data collection and assurance of privacy to the end-users. Blockchain can be effectively used in collecting significant pieces of information with respect to

examination of the patient and be able to screen the movement of the patients to make sure that they are adhering to the SOPs of COVID-19, taking social distancing measures seriously. Nonetheless, protection of the patient's identity was essential all this while. Since there is no central authority to speak of with blockchain technology, users have complete control over the data they send and receive through the blockchain network. Blockchain technology may be used by healthcare organizations while maintaining a strict emphasis on safeguarding patient privacy and identification at all costs. A similar framework was introduced in Europe, where blockchain technology was used in COVID-19 tracing by making use of Bluetooth. On the other, a German company has made use of a blockchain framework using cellphones, while making sure that the client's security is not violated in any manner [29].

3.4.3. Managing Supply Chain.
With the pandemic hitting the entire world, there were some significant interruptions across supply chains in all different sectors. The primary factors that were responsible for these interruptions were the shutting down of factories and improper or inadequate safety and hygiene measures followed in the facilities altogether. There was also an enormous demand for PPE kits and other important medical supplies to support the production and continuation of operations in corporations everywhere. Due to the lengthy supply chain, it was resulting

in excess obscurity that made it more difficult to track, calculate, and plan the supply processes altogether. Blockchain also ensures that the network is transparent and the transactions are securely broken down for further processing. As a result, large numbers of blockchain arrangements were incorporated into the supply chain during the pandemic. The advantages were substantial, it not only made it fairly quick to handle and process but also reduce the overall costs, low risk of operations, and faster settlement for all included within the network. There is a blockchain platform by the name VeChain, that is, ensuring a sustainable yet transparent supply of new KN95 masks from China while working with numerous production facilities and different corporations.

3.4.4. Outbreak Tracking. Amidst the outbreak of COVID-19, it was the most difficult task for the government to identify the infected patients and then get them to be secluded from the rest to gain control over the wide spread of the virus altogether. Blockchain removes the potential need for any kind of third party since it is highly decentralized that can potentially reduce the occurrence of data modification to a major extent. It is due to the use of blockchain technology that has potentially enhanced the reliability of the information available on the pandemic for the population in general. Evidently, fraudulent data and false information not only contribute to the enormous damage made to the economy as a whole but also end up affecting the psychological state of the general population altogether. Therefore, a blockchain database is highly reliable in such a vulnerable state of events, where it probably makes it more difficult to modify or make the data nontraceable on the whole.

Because of the accuracy, reliability, and transparency that blockchain technology provides, it is an ideal tracking platform. Governments can better update their databases as well with respect to coronavirus' current status and further improve the planning and management of the outbreak as well. This technology can also help in identifying the territories that are completely no-go zones and further tracking down the spread of the infection altogether. Hash-Log, a public database maintained by the prominent IT company Acer, provides information on the propagation of infections and allows users to follow the patterns over a certain period of time. For clinical studies, the Acer dashboard additionally makes use of data from the Centers for Disease Control and Prevention and the World Health Organization (WHO).

4. Conclusion

COVID-19 pandemic has affected numerous sectors of day-to--to-day life like healthcare, education, politics, and the economy to mention a few. Blockchain technology plays a major role in the management and development of preventive measures. Because the epidemic has aggravated the crisis of conscience in the administrative and medical centers as a result of the collapse to supply of timely and accurate facts about the current predicament. Hence, it

leads to the requirements of Covid- 19 data handling to provide all benefits to the user in which they suffered due to misguidance, and therefore to acquire certain features like data security, data integrity, real-time data handling data management, etc. COVID-19 data handling by the employment of blockchain framework plays a significant role. Hence, in this paper, the key features of blockchain technology have the potential to support the implementation of many use cases such as contact tracing, efficient data collection, secure data storage, and effective data analysis. Comparing Blockchain data handling systems to centralized database systems, blockchain provides the potential to improve COVID-19 Data handling to a major extent.

Before the blockchain can be fully implemented to combat COVID-19 and improve COVID-19 preventive measures, a number of problems, including legal issues, latency issues, resource usage issues, and widespread implementation issues, must be handled. We put out a number of potential options for enhancing blockchain adoption in the future in light of the COVID-19 outbreak. By offering a secure framework for capturing, maintaining, and sending sensitive data, blockchain applications may be utilized to decrease network latency. Furthermore, COVID-19 management will be made possible by the recent integration of blockchain with other cutting-edge technologies including artificial intelligence, big data, and cloud computing. It is also suggested that blockchain is being coupled with other current technologies to give an adequate performance in resolving issues related to the COVID-19 pandemic to build a strong healthcare infrastructure. Furthermore, by utilizing new cutting-edge and distinctive security methodologies, the corresponding security risk regarding blockchain technology can be addressed. To address data vulnerabilities in blockchain consensus protocols, security defense solutions are gathered under this heading. A recipient-oriented transaction plan that uses the phrases "stealth address and master node" to authorize the transaction before adding it to the block is presented as a solution to the problem of double spending in blockchain. In this case, both the transaction's originator and the recipient actively take part in the broadcast process's transaction verification. The patient's data can be made more secure and protected overall by implementing this feature in the healthcare industry. Finally, the blockchain-based application needs to be upgraded to outperform in several technological aspects with the aim of employing the blockchain as a model alternative for the crisis in the healthcare domain, akin to the COVID-19 pandemic. Scalability and lightweight blockchain implementation, for instance, are essential in the healthcare industry to minimize data verification and delay. To minimize processing time and enable high-speed data sharing, effective mining techniques may be utilized in conjunction with shorter verification operation times, leading to robust data analysis in a setting identical to the COVID-19 environment.

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. Alam, F. Ahmad Reegu, S. M. Daud, and M. Shuaib, *Blockchain-Based Electronic Health Record System for Efficient COVID-19 Pandemic Management*, MDPI AG, Basel, Switzerland, 2021.

[2] M. M. Khubrani and S. Alam, "A detailed review of blockchain-based applications for protection against pandemic like COVID-19," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 4, p. 1185, 2021.

[3] T. Kumar, V. Ramani, I. Ahmad, A. Braeken, E. Harjula, and M. Ylianttila, "Blockchain utilization in healthcare: key requirements and challenges," in *Proceedings of the 2018 IEEE 20th International Conference on E-Health Networking, Applications and Services (Healthcom)*, September 2018.

[4] J. Sengupta, S. Ruj, and S. Das Bit, "A comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT," *Journal of Network and Computer Applications*, vol. 149, Article ID 102481, 2020.

[5] G.-T. Nguyen and K. Kim, "A survey about consensus algorithms used in blockchain," *Journal of Information Processing Systems*, vol. 14, no. 1, pp. 101–128, 2018.

[6] S. M. H. Bamakan, A. Motavali, and A. Babaei Bondarti, "A survey of blockchain consensus algorithms performance evaluation criteria," *Expert Systems with Applications*, vol. 154, Article ID 113385, 2020.

[7] L. Lenert and B. Y. McSwain, "Balancing health privacy, health information exchange, and research in the context of the COVID-19 pandemic," *Journal of the American Medical Informatics Association*, vol. 27, no. 6, pp. 963–966, 2020.

[8] J. Wang, C. Xia, A. Sharma, G. S. Gaba, and M. Shabaz, "Chest CT findings and differential diagnosis of mycoplasma pneumoniae pneumonia and mycoplasma pneumoniae combined with streptococcal pneumonia in children," *Journal of Healthcare Engineering*, vol. 2021, Article ID 8085530, 10 pages, 2021.

[9] D. Arias-Garzón, J. A. Alzate-Grisales, S. Orozco-Arias et al., "COVID-19 detection in X-ray images using convolutional neural networks," *Machine Learning with Applications*, vol. 6, Article ID 100138, 2021.

[10] S. Guefrechi, M. B. Jabra, A. Ammar, A. Koubaa, and H. Hamam, "Deep learning based detection of COVID-19 from chest X-ray images," *Multimedia Tools and Applications*, vol. 80, no. 21–23, pp. 31803–31820, 2021.

[11] G. Sargsyan, N. Castellon, R. Binnendijk, and P. Cozijnsen, "Blockchain Security by Design Framework for Trust and Adoption in IoT Environment," in *Proceedings of the 2019 IEEE World Congress on Services (SERVICES)*, July 2019.

[12] B. Xie, D. He, T. Mercer et al., "Global health crises are also information crises: a call to action," *Journal of the Association for Information Science and Technology*, vol. 71, no. 12, pp. 1419–1423, 2020.

[13] J. Pan, J. Wang, A. Hester, I. Alqerm, Y. Liu, and Y. Zhao, "EdgeChain: an edge-IoT framework and prototype based on blockchain and smart contracts," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4719–4732, 2019.

[14] F. M. Enescu, N. Bizon, A. Cirstea, and C. Stirbu, "Blockchain technology applied in health the study of blockchain application in the health system," in *Proceedings of the 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, June 2018.

[15] C. J. Galvin, L. Fernandez-Luque, and Y.-C. J. Li, "Accelerating the global response against the exponentially growing COVID-19 outbreak through decent data sharing," *Diagnostic Microbiology and Infectious Disease*, vol. 101, no. 2, Article ID 115070, 2021.

[16] H. Wu, Y. Shang, L. Wang, L. Shi, K. Jiang, and J. Dong, "A patient-centric interoperable framework for health information exchange via blockchain," in *Proceedings of the 2019 2nd International Conference on Blockchain Technology and Applications. ICBTA 2019: 2019 2nd International Conference on Blockchain Technology and Applications*, December 2019.

[17] I. Abu-elezz, A. Hassan, A. Nazeemudeen, M. Househ, and A. Abd-alrazaq, "The benefits and threats of blockchain technology in healthcare: a scoping review," *International Journal of Medical Informatics*, vol. 142, Article ID 104246, 2020.

[18] F. Ajaz, M. Naseem, S. Sharma, M. Shabaz, and G. Dhiman, "COVID-19: challenges and its technological solutions using IoT," *Current Medical Imaging Formerly: Current Medical Imaging Reviews*, vol. 17, 2021.

[19] H. Hasanova, U. Baek, M. Shin, K. Cho, and M.-S. Kim, "A survey on blockchain cybersecurity vulnerabilities and possible countermeasures," *International Journal of Network Management*, vol. 29, no. 2, Article ID e2060, 2019.

[20] Z. Zheng, S. Xie, H.-N. Dai et al., "An overview on smart contracts: challenges, advances and platforms," *Future Generation Computer Systems*, vol. 105, pp. 475–491, 2020.

[21] H. R. Hasan, K. Salah, R. Jayaraman, I. Yaqoob, M. Omar, and S. Ellahham, "COVID-19 contact tracing using blockchain," *IEEE Access*, vol. 9, Article ID 62956, 2021.

[22] A. Gupta and L. K. Awasthi, "Peer-to-peer networks and computation: current trends and future perspectives," *Computing and Informatics*, vol. 30, no. 3, pp. 559–594, 2011.

[23] D. Lee and N. Park, "Blockchain based privacy preserving multimedia intelligent video surveillance using secure Merkle tree," *Multimedia Tools and Applications*, vol. 80, no. 26–27, pp. 34517–34534, 2020.

[24] A. Prashanth Joshi, M. Han, and Y. Wang, "A survey on security and privacy issues of blockchain technology," *Mathematical Foundations of Computing*, vol. 1, no. 2, pp. 121–147, 2018.

[25] A. Gupta and L. K. Awasthi, "Peer enterprises: a viable alternative to Cloud computing?" in *Proceedings of the 2009 IEEE International Conference on Internet Multimedia Services Architecture and Applications (IMSAA)*, Bangalore, India, December 2009.

[26] S. Chaudhury, N. Shelke, K. Sau, B. Prasanalakshmi, and M. Shabaz, "A novel approach to classifying breast cancer histopathology biopsy images using bilateral knowledge distillation and label smoothing regularization,"

Computational and Mathematical Methods in Medicine, vol. 2021, Article ID 4019358, 11 pages, 2021.

[27] H. Jin, C. Xu, Y. Luo, P. Li, Y. Cao, and J. Mathew, "Toward secure, privacy-preserving, and interoperable medical data sharing via blockchain," in *Proceedings of the 2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS). 2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*, December 2019.

[28] A. H. Mayer, C. A. da Costa, and R. da R. Righi, "Electronic health records in a Blockchain: a systematic review," *Health Informatics Journal*, vol. 26, no. 2, pp. 1273–1288, 2019.

[29] H. Rathore, A. Mohamed, and M. Guizani, "Blockchain applications for healthcare," *Energy Efficiency of Medical Devices and Healthcare Applications*, vol. 2020, pp. 153–166, 2020.

Hindawi

Research Article

Design and Implementation of Financial Service and Management Platform considering Support Vector Machine Algorithm

Lei Tian [ID]

School of Economics and Management, Chifeng University, Inner Mongolia, Chifeng, China

Correspondence should be addressed to Lei Tian; 1551140013@xzyz.edu.cn

With the rapid economic development, the financial industry has quietly become the leader of industries, the core and lifeblood of promoting economic development. At the same time, various financial services and management platforms emerge one after another. However, the emergence of financial services and management platforms cannot effectively alleviate the current financial crisis. In the face of increasingly complex financial risks, traditional financial service and management platforms cannot achieve effective information sharing, which leads to continued low service and management efficiency and frequent financial risk problems. Support vector machine is a data classification algorithm based on supervision, which can realize data sharing and improve the efficiency of data processing. The article firstly readjusted the underlying architecture of the financial service and management platform to break through the barriers of data interaction. Then on this basis, the article further combines the support vector machine algorithm and extends it from binary data classification to multivariate classification. Finally, the paper redesigns the financial service and management platform considering support vector machines. After a series of experiments, it can be found that the financial service and management platform based on the support vector machine algorithm can reduce the financial risk by 17.2%, improve the financial service level by 30.2%, and improve the financial comprehensive service level by 45.2%. At the same time, thanks to information sharing and interaction, the financial service and management platform can effectively predict financial risks, and the accuracy of the prediction basically reaches 78.9%. This shows that a financial service and management platform that takes into account the support vector machine algorithm can effectively prevent financial risks and improve the efficiency of financial services and management.

1. Introduction

With the rapid development of the Internet, the financial industry has gradually penetrated into every corner of life and brought a great impact on the traditional financial system. In order to cope with the rapidly developing economy and avoid major financial turmoil, many banks and core enterprises have established their own financial service and management platforms. However, for a long time, the quality of financial services has not been improved because the information between financial institutions and enterprises has not been communicated with each other, and many enterprises have suffered greatly. At the same time, due to the blockage of information, the comprehensive efficiency of financial services and management is low, and financial risks occur frequently.

The support vector machine algorithm can fully mine the data between financial institutions and enterprises, promote information sharing between financial services and management platforms, and improve the efficiency of financial services and management. At the same time, by integrating financial resources, relevant institutions can use the platform to improve their ability to monitor, assess, and prevent risks in financial institutions. On this basis, financial institutions can further enhance their authority. At the same time, by building a financial service and management platform that takes into account the support vector machine algorithm, it can further accelerate the innovation of financial service products, improve the relevant system of financial management, enhance the ability of the financial industry to resist risks, and promote the stable and long-term development of the financial field.

After a series of experiments, we knew that the financial service and management platform based on support vector machine can greatly improve the level of financial services. Before the support vector machine is not taken into account, financial services cannot synthesize all information, and the core enterprises provide financial services with low efficiency and single products, and the service level is as low as 11.2%. After focusing on support vector machines, the service level of the platform increased by 30.2% year-on-year, and the comprehensive average service level of the platform increased to 45.2%. In the process of financial risk prediction and assessment, the Bayesian algorithm can effectively predict financial risks caused by supervision with an accuracy of 71.2%. The risk prediction method based on neural network can effectively predict the risk of financial investment with an accuracy of 71.9%. In contrast, the financial service and management platform based on support vector machine can better predict various business risks in the financial field, and its prediction accuracy can reach up to 78.9%. This fully shows that the financial service and management platform based on support vector machine can effectively improve the efficiency of financial services and financial management level and can predict financial risks in a timely manner.

2. Related Work

With the continuous development of the financial industry, the problems encountered in financial services and management are becoming more and more serious. Many scholars have discussed this. Chang-Ho pointed out that as the global economy continues to slump, companies with low liquidity and low profitability will face a recession, which is expected to increase the risk of bankruptcy. Liquidity and profitability are closely related in the analysis of operating performance. Therefore, he used a standardized correlation matrix to establish a simple financial service platform to analyze the financial factors that affect the liquidity and profitability of financial enterprises [1]. Dhiman pointed out that the financial and banking industry has made great progress in technology, so he pointed out that it is necessary to deeply study the successful template of digital transformation and quantify it. Taking financial enterprises in the digital age as the research object, he focused on the help of digital technology in the financial field to human decision-making and emphatically analyzed the impact of financial management methods on the economic development of enterprises [2]. Piatti-Fünfkirchen and Schneider pointed out that the mobile financial service platform is an innovation of the traditional financial platform and has strategic significance for the cross-border development of the global financial industry. In the research process, they studied the mobile financial service platform as a complex ecosystem and put forward the development direction and related suggestions of the mobile financial service platform [3]. Omigie et al. analyzed the motivation of financial product innovation in order to explore the risk of banking financial services under the background of informationization. Therefore, they established a new financial product service

platform with spillover effects, different competitive strategies, and risk constraints. On the basis of analyzing the risk characteristics of financial product innovation, they discussed the risk transmission mechanism of the financial product innovation chain and proposed a new model of internal risk management and external risk supervision of bank financial product innovation [4]. Ali and Isak's research examined the impact of financial management platforms on financial management mechanisms. During the research process, they used explanatory and descriptive research designs, using stratified sampling techniques to select a sample of 145 respondents [5]. Although the above scholars have analyzed the problems of financial services and management from different levels, they have not found the essence of the problem.

Support vector machines are often used in data classification and mining, which can accurately find problems in financial services and management, and many scholars have also conducted research on this. Guo et al. pointed out that smart hospitals are considered a promising technology and platform that can greatly improve healthcare in the future. In order to further strengthen the platform construction, they analyzed two important parameters of the signal pulse through the support vector machine (SVM) algorithm and accurately divided all cellular events into two different subpopulations. Then the microfluidic sensor is combined with the SVM algorithm to reconstruct the sensor network system of the smart hospital [6]. Hao et al. proposed a comprehensive SVM classification algorithm for the imbalance of training data set. First, they performed unsupervised clustering on the imbalanced training set, then segmented the data set with SVMs to precisely control the local features of the data set, and finally solved the data imbalance [7]. Shyamala et al. pointed out that in the gas pipeline system, the safe operation of the gas pressure regulator determines the stability of the fuel gas supply. Therefore, according to the operating conditions of the medium and low pressure gas regulators in the SCADA system, they proposed a new method for the safety precaution of gas regulators based on SVMs. This method takes the gas pressure regulator outlet pressure data as the input variable of the SVM model and takes the fault type and degree as the output variable, which can effectively improve the prevention accuracy and save a lot of manpower and material resources [8]. Sontayasara et al. pointed out that bearing failure is the main cause of motor and generator failure, so they proposed to use SVM algorithm for early detection and classification of bearing failure. By using his proposed method, bearing failures can be detected at an early stage, giving machine operators time to take preventive measures before large-scale failures occur [9]. Zhu pointed out that the precise analysis of the data is essential to improve the reliability of the diagnostic method. Therefore, based on solid-state micropore microfluidic impedance cytometer, he used the SVM algorithm to detect and enumerate cancer cells in red blood cells [10]. The above experts and scholars have fully exploited the advantages of SVMs, but they have not extended the algorithm to the financial field, and their research on the SVM algorithm is not very thorough.

3. Support Vector Machine Algorithms and Financial Services and Management

3.1. Financial Services and Management. Finance is an economic activity in a market environment, and it represents the entire process of monetary and capital financing. With the rapid economic development, the financial industry has gradually become a barometer of economic development. Therefore, vigorously developing financial services and management is an important measure to promote economic development [11]. Under the influence of the global economic crisis, many industries around the world have not completely emerged from the shadow of the financial crisis. Among them, financial enterprises are particularly affected. After experiencing many economic crises, people gradually realized the importance of strengthening financial management and innovating financial services. Therefore, the majority of financial enterprises and institutions slowly explored and established their own financial services and management systems. In the traditional financial industry, financial services and financial management are actually two completely different concepts. Therefore, the article will explore its existing problems from two levels.

"Financial services" is a general term for financial institutions to use financial instruments to provide information or other services required for investment. In the process of financial services, financial institutions usually provide necessary information for participants and customers, and then customers carry out a series of financial activities based on the relevant information. However, in the process of information exchange, there are obvious information barriers between financial institutions and customers and participants. On the one hand, most of the information is in the hands of financial institutions and financial core enterprises, and customers cannot directly obtain first-hand information and data [12]. On the other hand, the data and related information of customers and event participants using information to engage in financial activities also remain at the customer level and cannot be finally aggregated into financial institutions. Based on this, the information between financial institutions and customers is completely separated, which greatly increases the risk of financial activities, and also brings bad experience to customers. In general, the backward concept of financial services and the singleness of financial service products are the disadvantages of the old financial service model. It is precisely because of the above disadvantages of traditional financial services that it is of great significance for the majority of financial institutions and core enterprises to continuously enhance the awareness of financial services and improve the level of financial services.

Financial management refers to the whole process in which financial institutions and enterprises supervise and coordinate currency circulation and credit activities in accordance with certain financial activity laws. Under the current economic environment, the majority of financial institutions focus their financial management on credit management. The emphasis on credit management is determined by both historical factors and reality. On the one hand, financial crises in history are often caused by credit crises, so strengthening credit management can prevent problems before they occur. On the other hand, in the current economic environment, in the face of business expansion and performance, a large number of illegal loans have appeared in the majority of financial institutions. But only focusing on one area of management can lead to losing all the game [13]. At the same time, the existing financial management methods cannot adapt to the financial model in the new era. Therefore, the continuous innovation of financial management methods and means is an important measure to achieve the sustainable development of the financial industry.

At the level of financial services and financial management, many financial institutions have not fully realized the importance of information exchange and the necessity of establishing a sound management system. At the same time, data and information about the financial sector are basically only circulated among a few core companies. Under such circumstances, the financial services and management systems established by financial institutions and enterprises have inherent drawbacks. Among them, the financial management and service system under the traditional model is shown in Figure 1.

In the traditional financial management and service system, financial institutions and enterprises often focus on the circulation of funds, so funds are often placed in the center of the financial service and management system [14]. In this case, other participants in financial activities often form separate contacts, which greatly reduces the efficiency of service and management and increases the risk of financial activities. In addition, there is a lack of obvious supervision links in the traditional financial management and service system, which has laid hidden dangers for the financial activities of enterprises. Moreover, by observing the architecture, it can be found that the traditional financial model has significant instability, so it is imminent to optimize and upgrade the system.

3.2. Financial Services and Management Platforms Focusing on SVM. In the above, the feasibility of combining SVMs with financial services and management platforms is analyzed from the theoretical level. And the utility of the financial service and management platform focusing on SVMs is explored from multiple perspectives. However, there are differences between theory and reality, and the financial cycle and its fluctuations cannot be measured theoretically, so empirical research will be a strong guarantee to test the utility of the platform. At the same time, in order to ensure the accuracy of the data, some data were obtained from the Wande database and the Tai'an database, and the statistical results of the obtained sample data sets are shown in Table 1.

Table 1 shows that the sample data set mainly includes factors such as financial cycle, financial volatility, and economic growth level. Among them, the smaller the financial volatility, the faster the economic growth rate. In order to obtain general laws and methods from general data,
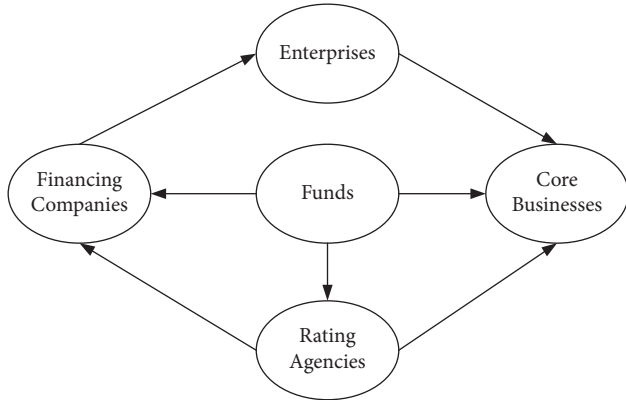
Figure 1: Financial management and service system under the traditional model.

Table 1: Sample data set statistics.

Variables	Average value	Variance	Minimum value	Maximum value
Cycle	0.768	0.611	−0.098	2.115
Vol	0.091	0.083	0.006	0.732
InGDP	3.642	0.381	1.002	4.216
Capital	0.492	0.151	0.209	1.395

the above data sets are added to the training of SVMs. At the same time, in order to further test the correctness of the data classification results of SVMs, this paper uses three different structures of matrices to evaluate the results. The specific empirical results are shown in Table 2.

Table 2 shows that by testing most of the factors, the classification results of the SVM are basically accurate, which lays the foundation for the further experiments of the article. Among them, the financial cycle coefficient remained positive during the three tests and passed the 1% significance level test. This shows that changes in the financial cycle are consistent with the direction of changes in economic growth. In addition, the coefficient of financial volatility is significantly negative in multiple tests, which indicates that there is a negative correlation between it and economic growth.

After the above inspection and evaluation, it is reasonable to believe that the classification and decision-making based on support vectors are accurate. Therefore, people will further examine its role in financial services and management platforms. LR test is an indicator that can best reflect the authenticity of the data, which can reflect the sensitivity and specificity of the model at the same time, so people will use this test method to test the data on the above platforms. Among them, the LR test results of the financial platform based on SVM are shown in Table 3.

Table 3 shows that in the face of different financial risks, the test results of financial platforms based on SVMs are not the same. Among them, technical risks and operational risks are unavoidable during the operation of the platform. Therefore, when dealing with these two types of risks, the LR test value of the platform is relatively high, and the maximum lag value is 11.654. In addition, in the face of liquidity risk and credit risk, the platform's LR test value is low, with a

Table 2: Empirical results under multiple factor assessment.

Variables	Adjacency matrix	Diagonal matrix	Inverse matrix
Models	SVM	SVM	SVM
Cycle	0.231***	0.038***	0.019***
Vol	−0.035	−0.029**	−0.034***
InDA	0.0139	0.0115	0.0451*
Capital	0.031*	0.019	0.214

Note. *, **, and ***represent data at 10%, 5%, and 1% confidence levels, respectively.

Table 3: Support vector machine-based financial platform LR test.

	Credit risk	Liquidity risk	Operational risk	Technical risk
Hysteresis	11.058	5.214	9.021	10.654
Error	12.021	8.399	11.654	9.687
Regulation	16.215	10.005	18.001	2.134

minimum of 5.214. This shows that in dealing with financial risks, the financial platform of the SVM can basically pass the test, and the model selection meets the requirements.

3.3. Support Vector Machines. SVM is often used for data classification and mining, so it can fully mine all the information in the economic environment and classify it to improve the efficiency of information utilization [15]. In the financial environment, information and data are necessary weapons for financial activities, and the emergence of SVMs has provided convenience for the majority of financial institutions. In addition, SVMs not only have excellent characteristics in classification problems, but also have good generalization ability in similar problems. Among them, the basic application areas of SVMs are shown in Figure 2.

With the optimization and upgrading of the SVM algorithm, it has achieved remarkable results in machine vision and natural language processing. In the information age, all production and business activities are inseparable from the support of data, and this is particularly prominent in the financial industry. In particular, financial services and management are financial activities built on data, so the combination of SVMs and financial services and management has a certain information base [16]. In this environment, the SVM can divide the financial market into different categories, then select different financial service and management cases as training data sets, and finally obtain the optimal decision plane and decision function. The schematic figure of the combination of SVM and financial services and management is shown in Figure 3.

In the process of combining technology and reality, SVM has shown excellent scalability, which has laid a solid foundation for the combination of SVM and financial services and management [17]. Among them, from the perspective of geometric structure, the combination of SVM and financial services and management shows strong stability, so this is conducive to fully exploiting and minimizing risks in financial activities. From the perspective of data
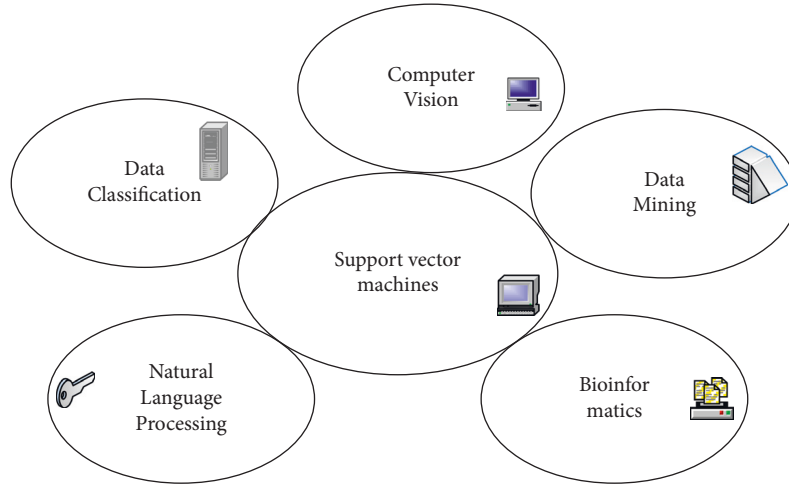
Figure 2: Basic application areas of SVMs.
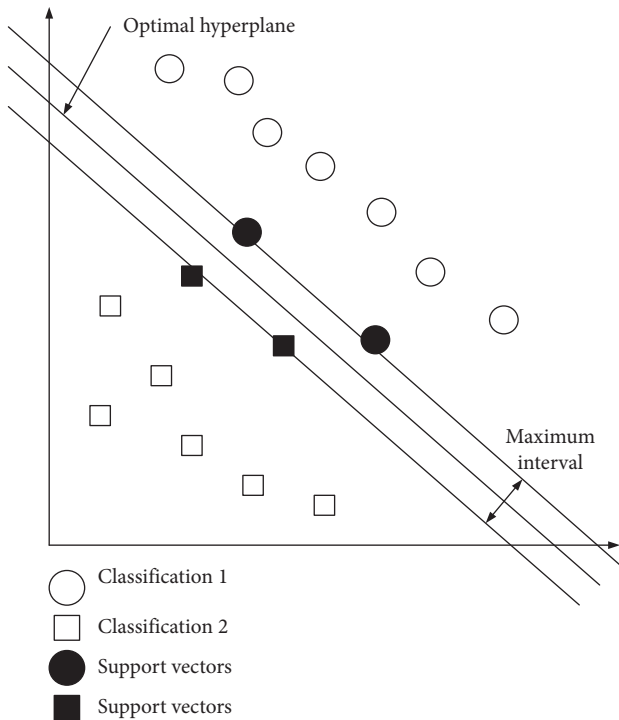
Figure 3: Schematic figure of the integration of SVMs with financial services and management.

distribution, the combination of SVMs and financial services and management takes into account both boundary conditions and internal conditions of the sample, so the combination of the two shows strong inclusiveness [18]. It is precisely because the combination of the two realizes the true meaning of learning from each other's strengths and complementing one's weaknesses, so it is particularly important to study them. In the above, the feasibility of combining SVMs with financial services and management is analyzed from the aspects of domain attributes and realistic conditions, and then it will be formally incorporated into the design of financial services and management platforms.

Assuming that there is a linearly separable data set T in the financial activity space, which represents the sample set in the training process, it can be mathematically expressed as

$$T = \{(a_1, b_1), (a_2, b_2), (a_m, b_m)\},$$

$$\gamma_m = b\left(\frac{\omega}{\omega} \bullet a_m + \frac{b}{\omega}\right). \tag{1}$$

Among them, a represents the feature vector, b represents the pointer, γ_m represents the straight-line distance between the points in the data set, and the positive or negative of ω represents the direction of the distance. People specify that positive values represent the positive direction and negative values represent the opposite direction. In this process, information about financial services and management is put into the data set.

Next, people define a hyperplane s; then, the minimum value of the distance between the sample point and the plane is calculated as

$$\varepsilon = \min_{i=1,2,3\bullet\bullet,M} \gamma_s. \tag{2}$$

In the above formula, the s plane is called the hyperplane, so the distance ε actually represents the distance from the support vector to the hyperplane. According to the above formulas, the hyperplane can be visualized as a financial market, then the original hyperplane distance problem can be expressed as the following capital circulation optimization problem.

$$y_i\left(\frac{\omega}{\omega} \bullet a_m + \frac{b}{\omega}\right) \geq \gamma. \tag{3}$$

In the above formula, what needs to be explored is the distance between the points in the data set and the plane, that is, the correlation between the financial activity information and the financial market. At the same time, it is expected to use this distance to maximize the segmentation of the data set, so as to find the most valuable information for financial activities. Therefore, for the simplicity of expression, (3) is calculated as follows:

$$y_i\left(\frac{\omega}{\omega\gamma}\bullet a_m + \frac{b}{\omega\gamma}\right) \geq 1, \tag{4}$$

$$\mu = \frac{\omega}{\omega\gamma},$$

$$b = \frac{b}{\omega\gamma}, \tag{5}$$

$$y_i\left(\mu\bullet a_m + b\right) \geq 1, (m = 1, 2, ..., M.$$

In the above formula, first, on the basis of (3), both sides are divided by γ, and then (4) is obtained. Then the scalars ω and γ are replaced by the vectors μ and b through a series of iterative calculations, thereby obtaining (7). After the above operations, the massive information is divided into different categories, and the same data are optimized.

The premise of the above optimization problem is a binary classification object, but in the process of actual financial activities, multiclassification problems are often encountered. Quadratic programming can transform multiclassification problems into binary classification, from which people can get

$$L(\omega, b, \alpha) = \frac{1}{2}\omega^2 - \sum_{i=1}^{M}\alpha_i\left(y_i\left(\mu\bullet a_m + b\right) - 1\right), \tag{6}$$

$$\theta(\omega) = \max_{\alpha_i \geq 0} L(\omega, b, \alpha). \tag{7}$$

Among them, L represents the Lagrangian function, and α represents the Lagrangian multiplier. Observing (6), it can be found that there are three variables in the newly constructed objective function, but there are only two pairs of existing constraints, so the optimal solution of the function cannot be solved only by the above formula.

Using Lagrange's duality, (6) can be reformulated as

$$\max_{\alpha_i \geq 0} \min_{\omega, b} L(\omega, b, \alpha) = c^*,$$

$$\omega = \sum_{i=1}^{M}\alpha_i y_i, \tag{8}$$

$$b = \sum_{i=1}^{M}\alpha_i y_i = 0.$$

In the above formula, the original optimal solution problem can be transformed into the problem of solving the maximum value. Similarly, for similar maximum segmentation problems, we can transform it into finding the maximum value.

By solving the maximum value, the optimal solution for the current situation can be obtained:

$$\alpha^* = \left(a_1^*, a_2^*, ..., a_m^*\right)^T,$$

$$0 \leq a^* \leq d, \tag{9}$$

$$a(x) = \text{sign}\left(\omega^*\bullet x + b\right).$$

In the above formula, α^* represents the optimal solution, and α represents the optimal classification decision function.

Through a series of calculations, people can finally use the data in financial activities to carry out financial services and get the most favorable decision-making plan.

In the above, the data set defined by people is linearly separable, but in practice there are very few completely linearly separable data in financial activities. To solve this problem, the concept of soft spacing is introduced.

$$\min_{\omega, b, \delta} \frac{1}{2}\omega^2 + d\sum_{i=1}^{m}\delta_i,$$

$$y\left(\omega\bullet a_m + b\right) \geq 1 - \delta_i, \tag{10}$$

$$\delta_i = \max\left(0, 1 - y\left(\omega\bullet a_m + b\right)\right).$$

Among them, δ represents a random variable, which allows some points in the data set to not meet the basic constraints, and d represents the penalty function: the larger the value, the greater the penalty for the classification result. Then people perform a separation operation on the hyperplane:

$$\omega^* \cdot x + b = 0. \tag{11}$$

Finally, for the above nonlinear problem, one can refer to the linear solution problem and solve it using the duality of Lagrangian. Then the optimal solution is obtained and the decision function is finally positioned to obtain the best solution for financial services and management. Through the above operations, it is also found that for the financial service and management data set T, none of the training data samples in the final solution model appear. This fully demonstrates the excellent characteristics of SVMs; that is, the final result is only related to support vectors, and it is also greatly adapted to financial services and management activities.

3.4. Design of Financial Service and Management Platform Focusing on Support Vector Machines. Reviewing the traditional financial service and management system, it is found that it has not formed a complete information channel and lacks effective supervision and management in the process of financial activities. Therefore, in order to maximize the effectiveness of the combination of SVMs and financial services and management, the traditional financial services and management systems have been optimized and upgraded. Then a financial service and management platform considering SVMs is designed. The financial service and management platform under the SVM algorithm is shown in Figure 4.

On the basis of the traditional financial service and management system, we use the SVM algorithm to reoptimize and adjust it. The following is the specific design and implementation process. First, in the original financial service and management system, people retained the basic core financial enterprises and financial structures in the financial service and management platform under the SVM algorithm [19]. In the financial market environment,
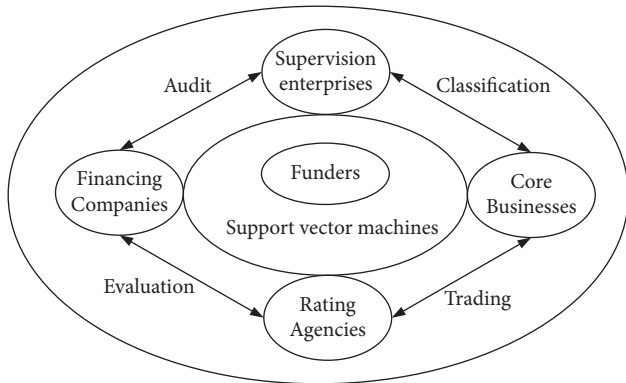
Figure 4: Financial service and management platform under the SVM algorithm.

financial enterprises and financial structures are the main body of the market environment. Therefore, no changes were made to the two during the design process. However, in view of the problem of noncommunication of information between various subjects, SVMs are used to bind the basic subjects in the financial market to data to achieve basic information communication. At the same time, in order to adapt it to the new financial service and management platform, it also provides data classification and tracking rights for these two subjects. Secondly, in the traditional financial service and management system, the financial supervision link is not prominent. Therefore, people have joined the supervision department in the financial service and management platform under the SVM algorithm. And in the design, people directly make it independent of the entire financial services and management links to achieve the autonomy of supervision [20]. Finally, in order to realize the exchange of information and promote the sharing of information among various subjects, a bidirectionally visible information module was added at the beginning of the platform design, which really improved the utilization rate of information from the bottom of the design. At the same time, in order to meet the financial market environment, the central position of funds in the financial service management system has been cancelled. Instead, it puts customers and investors at the center of the financial service and management platform to optimize the original system to the greatest extent and realize innovation in financial services and management. Moreover, the redesigned financial services and management platform achieves structural stability, which will aid its continued growth.

At the same time, in the implementation process of the financial service and management platform, the platform makes full use of SVMs to train and mine financial data. Then, according to the results, it can accurately profile financial institutions and customers, so as to help small- and medium-sized enterprises and small and micro enterprises to carry out financial activities [21]. In the process of profiling the enterprise, the relevant information of the enterprise will be circulated on the platform, so the enterprise applying for financial services can spend the least time to

maximize the benefits. At the same time, financial institutions can also publish financial products and services on the financial service and management platform, and enterprises that meet relevant conditions can apply directly on the platform to improve the efficiency and experience of financial services. For financial institutions, the application information of enterprises will be circulated directly on the platform, which avoids the leakage of information from multiple parties, and at the same time, it can further improve the efficiency of financial management. In this process, the financial service and management platform is customer-oriented, based on financial activity data, and uses SVMs and modern technology as means to build a multi-integrated, information-interoperable financial service and management model.

4. Deconstruction of Financial Services and Management Platform Results considering SVM

In the above, the financial service and management platform based on SVM has been empirically researched, but the theoretical test and pure data test cannot explain its effect in practical application. Therefore, the following article will explore the efficacy of the financial service and management platform one by one from the perspectives of the financial platform's service level, management accuracy, and risk assessment level. Among them, the financial service level taking into account the SVM is shown in Figure 5.

Figure 5 shows that the financial service and management platform based on SVM can greatly improve the level of financial services. Before the SVM is not taken into account, financial services cannot synthesize all information, the core enterprises provide financial services with low efficiency and single products, and the service level is as low as 11.2%. After focusing on SVMs, the service level of the platform increased by 30.2% year-on-year, and the comprehensive average service level of the platform increased to 45.2%.

The financial service level is only one part of the financial service and management platform, and it cannot describe the whole figure of the platform. Therefore, from the perspective of financial risk prediction and evaluation, the utility of financial service and management platform focusing on SVM in financial risk prediction is analyzed. The prediction results of different models on financial risk are shown in Figure 6.

In the financial field, the emergence of financial risks is often not without a trace. Figure 6 shows that in the process of financial risk prediction and assessment, the Bayesian algorithm can effectively predict financial risks caused by supervision, with an accuracy of 71.2%. The risk prediction method based on neural network can effectively predict the risk of financial investment with an accuracy of 71.9%. In contrast, the financial service and management platform based on SVM can better predict various business risks in the financial field, and its prediction accuracy can reach up to 78.9%.
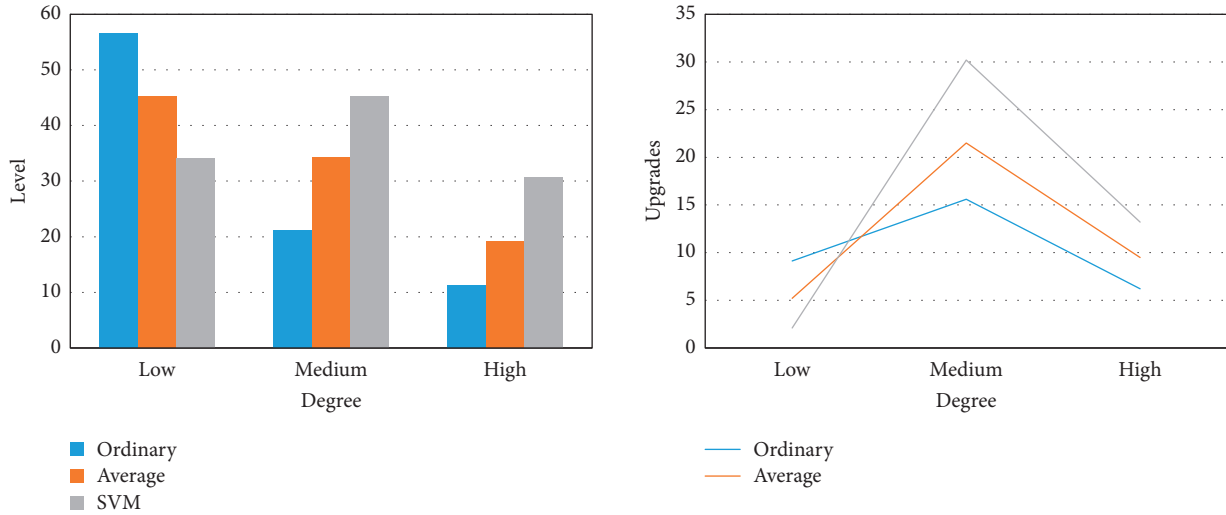
Figure 5: Financial services levels considering SVM.
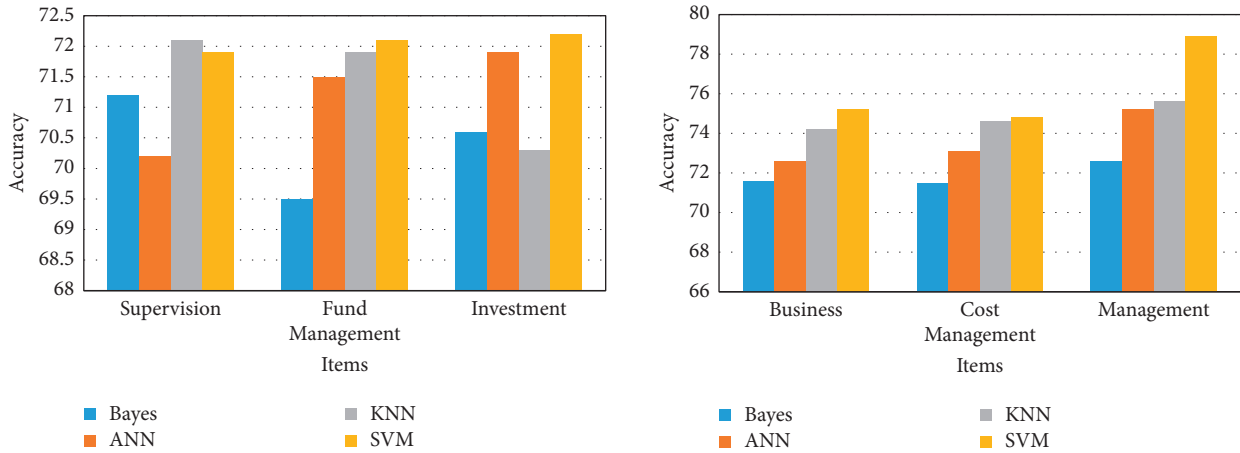
Figure 6: Prediction results of financial risk by different algorithms.

In the process of financial management and services, financial risks are everywhere. In the above, we have conducted experiments on the risk prediction ability of the platform. However, merely predicting risks cannot provide substantial help for subsequent financial investment behaviors. Therefore, it is necessary to evaluate the financial risks predicted by different platforms and conduct supervision and response according to the results of the evaluation, so as to minimize the risks. Among them, the risk assessment results of different platforms are shown in Figure 7.

Figure 7 shows that platforms based on different models have different risk factors. Among them, the platform risk coefficient based on the K series model is generally low, basically 18.3, and the platform risk based on the Bayesian model is relatively high, reaching 20.5%. In contrast, platforms based on SVMs have a lower risk factor of 17.2%.

The financial service and management platform is relatively comprehensive, so it is necessary to analyze its

efficiency from a comprehensive perspective. Among them, the comprehensive efficiency of the support vector machine-based financial service and management platform is shown in Figure 8.

The comprehensive efficiency of the financial platform is the efficiency of the financial sector serving the real economy. Figure 8 shows that the support vector machine-based financial service and management platform can amplify the role of data capital and promote the flow of funds across the industry. Among them, the financial service and management platform based on SVM can reduce the cost of trial and error, and its cost consumption is basically kept below 10.2%. In addition, the financial platform comprehensively utilizes multiparty information to realize the informatization and digitization of the platform, and its comprehensive proportion is 44.1% and 51.2%, respectively. This fully shows that the financial service and management platform based on SVM can comprehensively improve the efficiency of financial services and the level of financial management.
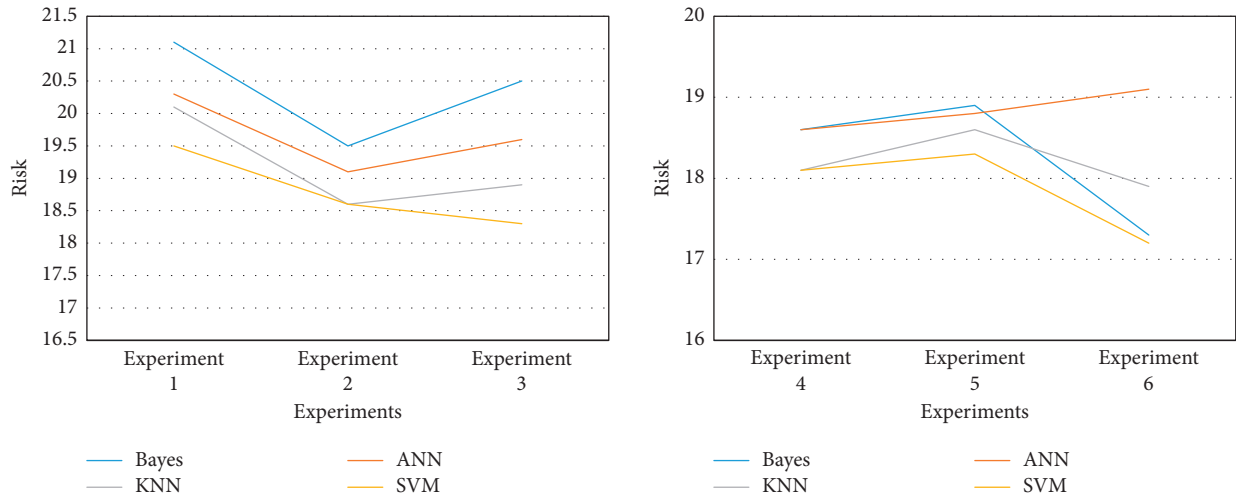
Figure 7: Risk assessment in financial management.
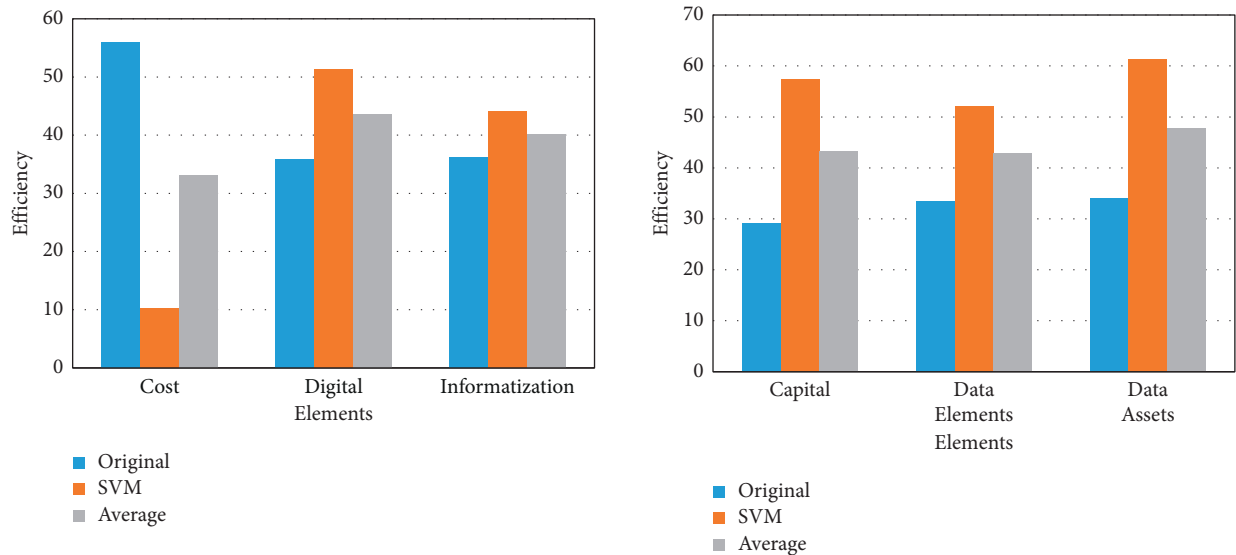
Figure 8: Comprehensive efficiency of financial platform based on SVM.

5. Conclusion

Financial risk is a major problem inevitably encountered in the development of the financial industry, and the breakthrough is to build a new financial service and management platform for information communication and business interaction. The article starts with the inherent problems of financial services and management and briefly analyzes the shortcomings and deficiencies of traditional financial services and management platforms. Then, the article optimizes and upgrades the method based on the SVM and integrates the SVM algorithm into the design of the financial service and management platform. Finally, the article focuses on the related methods and countermeasures of using SVMs to improve financial services and management. Experiments show that a financial service and management platform that considers support vectors can innovate financial services and

management and promote the sustainable development of the financial industry. However, due to time reasons, the financial service and management platform designed in this article still has some deficiencies in system security. In the future, the article will improve this aspect in the process of upgrading and optimizing the platform.

Data Availability

The data used to support the findings of this study can be obtained from the author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] A. Chang-Ho, "A study on the analysis of difference of management performance between financial services and information services industry," *The Society of Convergence Knowledge Transactions*, vol. 5, no. 2, pp. 41–47, 2017.

[2] R. Dhiman, "Identifying the key indicators of financial stability and financial development: a review of financial service sector," *Asian Journal of Management Science and Applications*, vol. 3, no. 4, pp. 302–320, 2018.

[3] M. Piatti-Fünfkirchen and P. Schneider, "From stumbling block to enabler: the role of public financial management in health service delivery in Tanzania and Zambia," *Health Systems & Reform*, vol. 4, no. 4, pp. 336–345, 2018.

[4] N. O. Omigie, H. Zo, J. J. Rho, and A. P. Ciganek, "Customer pre-adoption choice behavior for M-PESA mobile financial services Extending the theory of consumption values," *Industrial Management & Data Systems*, vol. 117, no. 5, pp. 910–926, 2017.

[5] A. Ali and A. O. Isak, "Financial management practices and financial performance of service companies in Somalia," *Research Journal of Finance and Accounting*, vol. 10, no. 4, pp. 59–68, 2019.

[6] J. Guo, Z. Chen, Y. L. Ban, and Y. Kang, "Precise enumeration of circulating tumor cells using support vector machine algorithm on a microfluidic sensor," *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 4, pp. 518–525, 2017.

[7] X. Hao, X. An, B. Wu, and S. He, "Application of a support vector machine algorithm to the safety precaution technique of medium-low pressure gas regulators," *Journal of Thermal Science*, vol. 27, no. 1, pp. 74–77, 2018.

[8] P. Shyamala, S. Mondal, and S. Chakraborty, "Numerical and experimental investigation for damage detection in FRP composite plates using support vector machine algorithm," *Structural Monitoring & Maintenance*, vol. 5, no. 2, pp. 243–260, 2018.

[9] T. Sontayasara, S. Jariyapongpaiboon, A. Promjun et al., "Twitter sentiment analysis of bangkok tourism during COVID-19 pandemic using support vector machine algorithm," *Journal of Disaster Research*, vol. 16, no. 1, pp. 24–30, 2021.

[10] M. Zhu, "Implementation of support-vector machine algorithm to develop a model for electronic commerce energy regulatory system," *Energy Reports*, vol. 7, no. 3, pp. 2703–2710, 2021.

[11] B. Gaye, D. Zhang, and A. Wulamu, "Improvement of support vector machine algorithm in big data background," *Mathematical Problems in Engineering*, vol. 2021, no. 1, pp. 1–9, 2021.

[12] L. Yang, G. Sun, A. Wang et al., "Predictive models of hypertensive disorders in pregnancy based on support vector machine algorithm," *Technology and Health Care*, vol. 28, no. 4, pp. 181–186, 2020.

[13] J. Wang, F. Yang, and X. Zhang, "Analysis of the influence mechanism of energy-related carbon emissions with a novel hybrid support vector machine algorithm in hebei, China," *Polish Journal of Environmental Studies*, vol. 28, no. 5, pp. 3475–3487, 2019.

[14] S. Mizrahi and Y. Minchuk, "Performance management in a decentralized setting: monitoring and gaming in the financial services industry," *Managerial and Decision Economics*, vol. 38, no. 6, pp. 742–753, 2017.

[15] Department of Business, "Ethical leadership, ethical selling, and financial services performances," *International Journal of Sustainable Economies Management*, vol. 6, no. 4, pp. 21–34, 2017.

[16] S. C. Misra and K. Doneria, "Application of cloud computing in financial services: an agent-oriented modelling approach," *Journal of Modelling in Management*, vol. 13, no. 4, pp. 994–1006, 2018.

[17] D. Makina, "Introduction to the financial services in Africa special issue," *African Journal of Economic and Management Studies*, vol. 8, no. 1, pp. 2–7, 2017.

[18] F. D. Zengul, R. Weech-Maldonado, B. Ozaydin, P. A. Patrician, and S. J. O'Connor, "Longitudinal analysis of high-technology medical services and hospital financial performance," *Health Care Management Review*, vol. 43, no. 1, pp. 2–11, 2018.

[19] W. W. Fenske, "Performance improvement and financial transition start with culture," *Frontiers of Health Services Management*, vol. 34, no. 2, pp. 25–30, 2017.

[20] M. Leyer, J. Strohhecker, and A. J. Kettenbohrer, "This business analytics tool looks nice, but. I am still happy without it - evidence from the financial services industry," *Journal of Service Management Research*, vol. 5, no. 1, pp. 20–35, 2021.

[21] T. Niemand, J. C. Rigtering, A. Kallmünzer, S. Kraus, and A. Maalaoui, "Digitalization in the financial industry: a contingency approach of entrepreneurial orientation and strategic vision on digitalization," *European Management Journal*, vol. 39, no. 3, pp. 317–326, 2021.

Review Article

A Comprehensive Study on Epidemiology Case Studies Using Computational Analysis

Xinjiang Lin,[1] **Shouping Chen** (ID)**,**[2] **and Amatul Bushra Akhi**[3]

[1]*Annual Ring Orthopaedic Hospital Group, Changsha, Hunan 410000, China*
[2]*Medical Management Center of Annulus Orthopaedic Hospital Group, Hunan 410000, China*
[3]*Department of CSE, Daffodil International University, Ashulia, Dhaka, Bangladesh*

Correspondence should be addressed to Shouping Chen; 1094426728@qq.com

Health-related issues and occurrences with regard to a particular population are the subject of an epidemiology study. This paper presents the results of a retrospective epidemiological investigation on 15922 hospitalized hand trauma patients from Central China between 2011 and 2020. Gender, age, onset season, injury mechanism, injury environment, injury location, and clinical characteristics are among the characteristics of the data gathered. The study is using computational analysis to draw inferences from the case studies collected in the databases of the hospitals. The types and characteristics of occupational injuries at home and outdoor are compared and analyzed. The purpose of the study is to present the findings from recent case studies for future reference and to recommend useful roles for the industrial sector in the care of patients with hand trauma in order to lower occupational harm. The injuries of preschool children are also analyzed. The incidence rate of hand injuries in infants has been increasing year by year which is directly related to the inefficient growth of children in rural areas. The data are collected from hospitals, then the data analytical tools are applied to draw conclusions. The suggested model is intelligently learned through the application of computational techniques, which are also used to suggest treatments to trauma victims. According to this study, males are more likely than females to sustain hand trauma; occupational injuries are more common than living injuries; males between the ages of 20 and 50 are at an increased risk of suffering an occupational injury. This study showed that the proportion of hand trauma in preschool children was higher (12.27%), and the 2-3-year-old group was the main injury target of preschool children (45.70%). The accidental injuries of newborns and young children can be reduced by government assistance, social support, and tighter monitoring.

1. Introduction

Human creation, life, and aesthetics all heavily rely on the hand. It is regarded as one of the most vital organs, second only to the face, and one of the body's most delicate portions. Thirty years ago, in [1] the authors analyzed 50,272 cases of hand trauma and found that 26% of them were injured at work, 34% were injured at home, and 15% were injured by participating in sports. The study's findings also demonstrate that various situations, including the workplace, the home, sports, and auto accidents, have their own unique internal laws that apply to hand injuries. Separate analysis and research are required. Twenty years ago, Jishuitan Hospital in China country reported that hand injuries accounted for

28.6% of emergency trauma patients [2]. In the past 10 years, research in China has shown that occupational hand trauma patients account for 50%–80% of the total number of hand trauma, and it is the most common occupational injury in the emergency department. Hand injuries account for about 37% of all occupational injuries. The incidence of hand trauma has significantly increased, primarily because of the quick growth of small private businesses and the relatively antiquated production equipment, as well as the lack of focus on safety facilities and safety education, unfamiliarity with the use of machines, and job changes. Factors such as temporary overtime and not wearing protective equipment have led to an increase in the incidence of occupational hand injuries year by year. We retrospectively analyzed the

epidemiological data of hand trauma in a significant orthopaedic hospital in Central China over the previous ten years. In order to study the epidemiological characteristics of hand trauma there, determine the regularity of its incidence, and provide a basis for its prevention and treatment.

Located in Central China, Changsha covers an area of 11,819 square kilometers and has a population of more than 10 million. It is an important provincial capital, a national historical and cultural city, a city with the best international image, and a leisure and entertainment city with a high happiness index. It is also a Chinese project. The capital of machinery, industrial manufacturing, construction, and handicrafts are also developed.

There are 30 large-scale tertiary hospitals in Changsha, and the medical and health services are well developed. The Annual Ring Orthopedic Hospital Group is a large-scale orthopaedic specialized hospital chain nationwide, with strong specialized technology, especially hand surgery technology. Therefore, more than 2,000 cases of hand trauma patients are received every year. In addition to the province as a whole, these patients originate from the city. From rural areas, road traffic, industry, and building sites, as well as from offices, service facilities, and schools. We believe that collecting and analyzing the epidemiological data of hand trauma patients in this hospital in the past 10 years is valuable for understanding and mastering the modern epidemiological history of the disease, the regularity of incidence, clinical characteristics, and assisting the government in formulating preventive measures to reduce the incidence.

The hands and digits of the human being occupy a major role in our professional life. Thus, not surprisingly, hand injuries are the most frequent bodily traumas sustained at work [3]. Even seemingly minor injuries can produce a devastating functional loss. Accurate, early diagnosis and definitive treatment are essential to minimize the functional loss and avoid the socioeconomic effects of protracted disability. Hand injuries due to occupational accidents or work-related hand injuries (WRHIs) lead to more serious consequences than any other organs in terms of both clinical courses and economic losses [4]. These injuries are as preventable as the other bodily traumas [5].

At least a million people are admitted to the emergency departments (ED) due to acute WRHI in the USA each year [6]. The incidence is particularly high in industries and jobs in which hand intensive work is done [5]. Hand trauma is a common surgical disease, and with the development of industrial machinery production, the incidence of acute occupational hand trauma is increasing [7]. Data show that acute occupational hand injuries account for 36.3% to 76.9% of emergency hand injuries [8], and Jishuitan Hospital in China reported that hand injuries account for 28.6% of trauma emergency patients [2]. In [9], the authors reported 732 cases of occupational hand trauma patients, the incidence of males was more than that of females. The epidemiological characteristics of occupational hand trauma patients reported by [10–12] are similar to the data in this study.

Among those with hand trauma, 76% were American men; 69% were British men; 59% were Danish men; 83% were Turkish men. In the Netherlands and Denmark, the age group between 10 and 14 year-old has the highest incidence. 44% of patients in Canada are 10–19 years old; 35% of patients in Turkey are under the age of 15 [13–16]. Among the occupational injuries in the statistics of the US labor department, hand trauma accounts for the largest proportion, accounting for 30% to 37% of all occupational injuries [17]. In [18], studies have shown that there is a statistical relationship between hand trauma and social deprivation, which varies with the type of injury.

Occupational injuries cannot be disregarded given China's tremendous advancement in science and technology, economy, and industrial manufacture. At present, China still lacks comprehensive and systematic big data analysis of hand trauma epidemiology data in the past ten years. This paper comprehensively and systematically analyzes and studies the distribution characteristics, injury mechanism, and clinical characteristics of hand trauma in different groups of people and provides a scientific basis for the government and enterprises to make decisions on the prevention and control of hand trauma.

Hand trauma is the most common trauma in people's production and life. According to literature reports, in emergency trauma cases, the proportion of hand trauma is between 15% and 28.6% [1, 19]. Reference [20]. So this study analyzes the percentage of hand trauma, patient's gender, age, onset season, injury mechanism, injury environment, injury site, and its clinical characteristics. The main highlights of the paper are as follows:

(1) Collect and analyze the data of 15,922 patients who met the criteria for hand trauma among inpatients from 2011 to 2020.

(2) To study the composition and distribution of hand trauma.

(3) To study the risk factors of hand trauma.

(4) Factors such as gender, age, the season in which the injury first occurs, the mechanism by which the damage occurred, the environment in which it occurred, the location of the injury, and its clinical characteristics are considered. The data are evaluated for preschoolers' injuries, life injuries, and occupational injuries, and compare them to domestic, Longitudinal, and horizontal comparative studies.

(5) Research on preventive measures of hand trauma is conducted.

Rest of the paper is organized as follows: in Section 1, background details of epidemiology hand trauma incidents and state of art of the epidemiology case studies are discussed. In Section 2, the proposed method of data collection, inclusion, and exclusion criteria of patients are discussed. In Section 3, result analysis and observations are discussed followed by a conclusion and references.

2. Proposed Method

2.1. Data Recording and Coding. Data from the Annual Ring Hospital Information System (HIS) is considered, which

included nine variables: admission date, month, and year; Sex; Age code; Age group; Injury caused; Product code and location code. The cause of injury was coded by International Classification of Diseases Ninth Edition (ICD-10).

2.2. Inclusion Criteria

(1) All injuries of the wrist joint and beyond; soft tissue injuries of upper limbs near wrist joints, such as nerve, tendon, blood vessel injuries, skin and soft tissue injuries in a large area, etc., but excluding injuries and burns of bones and joints within this range.

(2) The visit time was within 72 hours of injury.

(3) The medical records are complete.

2.3. Exclusion Criteria

(1) Chronic old hand injury.

(2) Hand diseases and deformities.

(3) Incomplete medical records.

2.4. Research Method.
There are 3 wards in the Hand Surgery Department of the Annual Ring Orthopedic Hospital, which consists of 15 physicians and 30 nurses, including 1 chief physician and 3 deputy chief physicians. Each ward has a part-time secretary. The director and secretary of the department are responsible for the quality of medical records and are required to fill in each medical record in a timely and complete manner, including the patient's basic information, injury, treatment, surgical methods, and recovery status and upload the HIS.

The previous 10-year medical records of hand trauma patients in the department were collected and sorted by the secretary of the department, and those who met the inclusion criteria were counted. In order to ensure the accuracy of the data, the hospital medical record room reviews the data compiled by each department and corrects the wrong information. The statistical contents include the following:

(1) General information (sex, age, occupation, etc.).

(2) Injury conditions (environment, nature, location, cause, etc.).

(3) Specific tissue damage (location and type, etc.).

(4) Treatment status (hospitalization time and outcome, etc.).

(5) Infants and young children morbidity, etc.

2.5. Statistical Analysis.
To estimate the frequency and percentage of all reported injuries, to describe the incidents, to specify the characteristics of injury brought on by consumer products, and to examine the risk factors, geographic locations, and seasonal patterns of injuries, descriptive epidemiological methods were used. SPSS 20.0 statistical software was used for analysis. The counting data was represented by the constituent ratio (%), and x^2 test was used for comparison between samples. Measurement data were expressed as $X \pm s$, and t-test was used for comparison between samples. $P < 0.05$ was considered statistically significant. Examine the injury's origin, location, season, type, and any possible correlations with age and sex. Age groups were analyzed according to six levels (0–6, 7–14, 15–18, 19–30, 31–65, and 66+).

3. Results and Analysis

3.1. Results

3.1.1. General Situation.
15922 inpatients with hand trauma were analyzed. Among them, 12897 cases were male (81.00%), 3025 cases were female (19.00%), and the ratio of male to female was 4.26 : 1. The incidence rate of the male was higher than that of females, and the difference was statistically significant ($x^2 = 6120.86$, $P = 0.001$). The oldest is 93 years old and the youngest is 5 days old. The mean age was (34.97 ± 18.23) years, including $(x = 35.71 \pm 17.50)$ years for males and $(x = 31.81 \pm 20.77)$ years for females. The gender distribution of each age group as shown in Figure 1 and is statistically significant. Among them, 49.21% were young and middle-aged patients aged from 31 to 65. Preschool children aged 0–6 years accounted for 12.27%.

Figures 2 and 3 depicts the annual and monthly distribution of in patients. The lowest is in February (5.63%) and the highest is in April (10.16%) every year.

3.1.2. Nature of Injury.
The main injuries were occupational injuries, with 10,866 cases (68.25%) and 5,056 cases (31.75%) of life injuries. Among the occupational injury patients, 9121 cases were male (83.94%), 1745 cases were female (16.06%), and the ratio of male to female was 5.23 : 1. The age was $(x = 40.93 \pm 16.28)$ years old, among which the age group was mainly from 20 to 50 years old, accounting for 74.18%.

There are two types of injured environments: indoor and outdoor. Indoor: factories, shopping malls, service places, leisure and entertainment places, office places, families, etc., are considered. For outdoor environment construction sites, roads, sports venues, fields, mountains, etc. According to their occupational injuries and life injuries, the patients in this group are different. Tables 1 and 2 shows the distribution of the injured environment and injured site.

3.1.3. Injured Area.
In this group, 8437 cases (52.99%) were injured in the left hand, 7244 cases (45.50%) in the right hand, and 241 cases (1.51%) in both hands. Among them, there were 6687 finger injuries, with a total of 7757 fingers. The distribution of finger injuries is shown in Figure 4. There were 162 finger injuries and 369 finger injuries with one hand. 3651 fingers were injured in the left finger and 3036
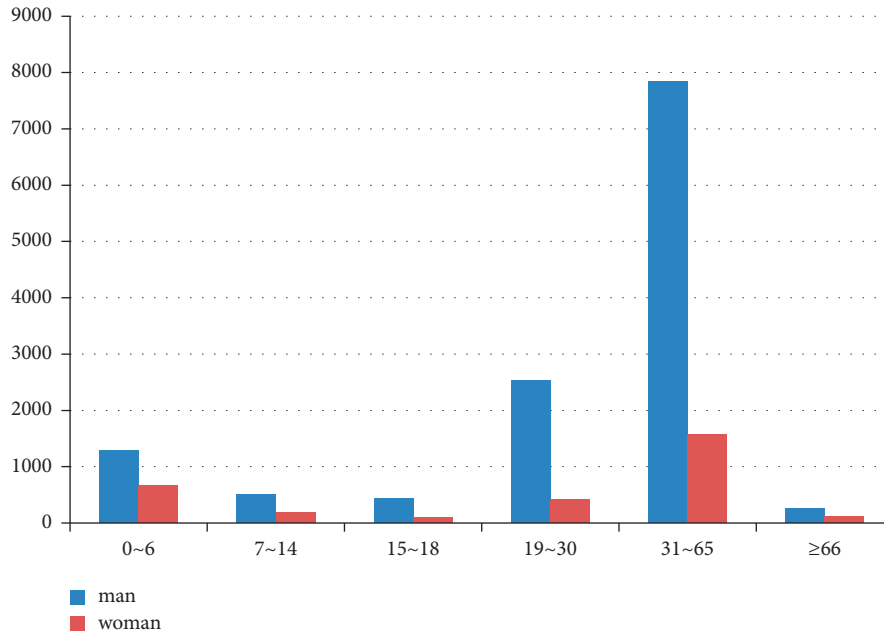
FIGURE 1: Distribution of age and sex. ($x2 = 22489.14$, $P = 0.001$).
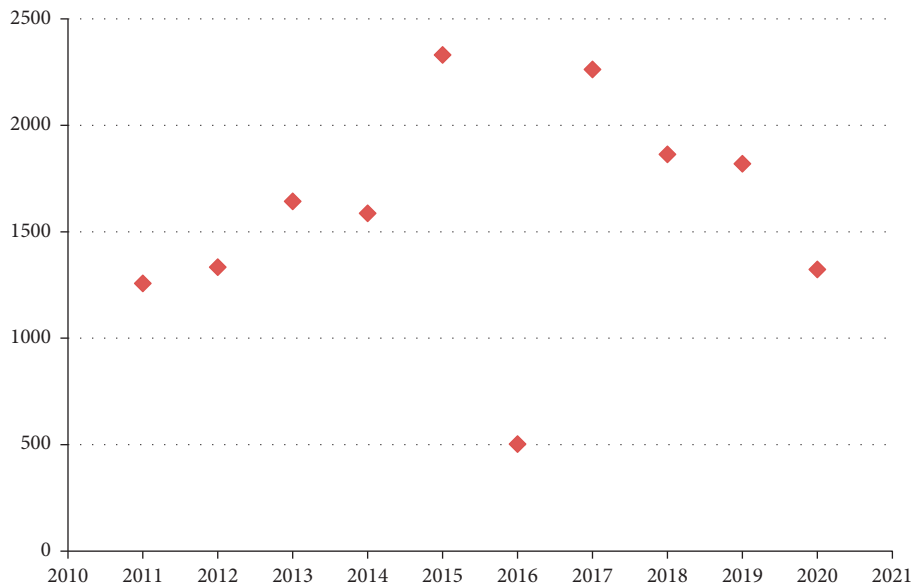
FIGURE 2: Distribution of the number of patients in each year ($x2 = 1607.94$, $P = 0.001$).

fingers were injured in the right finger. The ratio of left finger injury to right finger injury was 1.2 : 1.

There are different types of occupational injuries and life injuries, as shown in Table 3.

3.1.4. Morbidity in Preschool Children. There were 1954 preschool children aged 0 ~ 6 years (12.27%), including 1293 males and 661 females. Male: female is 1.96 : 1. Among them, there were 373 infants under 1-year-old (19.09%), 893 infants aged 2–3 years old (45.70%), and 688 children aged

4–6 years old (35.21%). The youngest is 6 days, because the baby is tied with ropes when wearing clothes, which causes strangulation and infection of the skin of both upper arms. The most common cause of injury was broken fingers (36.85%), as shown in Figure 5.

3.1.5. Hospital Stay. The length of hospital stay was ($x = 11.22 \pm 12.75$) days. 48.03% were hospitalized within 7 days. 8–14 days accounted for 31.38%; 15–21 days accounted for 12.15%; ≥22 days accounted for 8.44%, and the
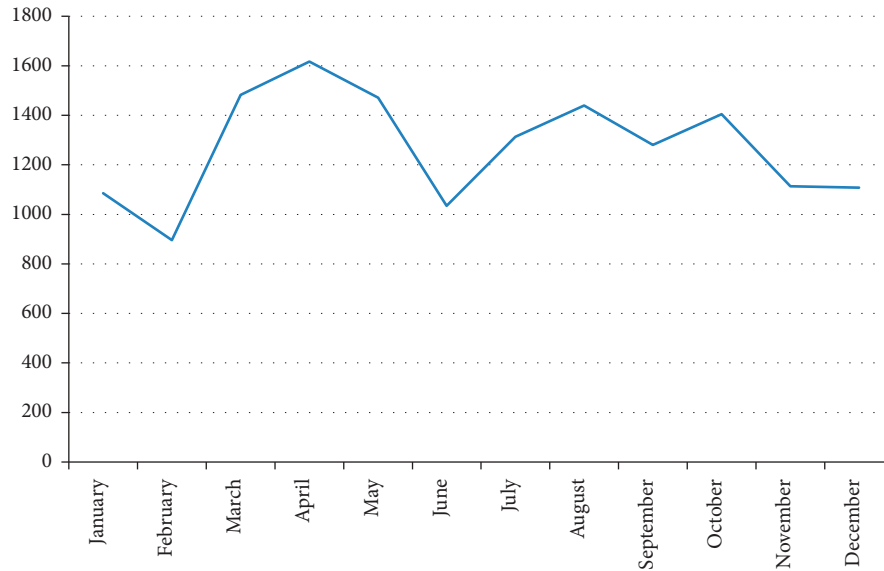
FIGURE 3: Distribution of morbidity in each month ($x2 = 555.05$, $P = 0.001$).

TABLE 1: Distribution of injured environment (%).

Classify	Interior	Outside	Total	Statistical analysis
Occupational	6209 (39.00)	4657 (29.25)	10866 (68.25)	$t = 6.999$, $p = 0.090$
Living injury	3660 (22.99)	1396 (8.77)	5056 (31.75)	$t = 2.231$, $p = 0.268$
Total	9869 (61.98)	6053 (38.02)	15922 (100)	

TABLE 2: Distribution of injured sites (%).

Classify	Construction site	Workshop	Road transport	Service place	Sports place	Office space	Leisure place	Field work	Other	Total	Statistical analysis
Occupational	3622	4242	622	1217	104	518	202	105	234	10866	$t = 2.273$, $P = 0.053$
Living injury	0	0	958	843	489	868	869	768	261	5056	$t = 4.347$, $P = 0.002$
Total	3622	4242	1580	2060	593	1386	1071	873	495	15922	
(%)	22.75	26.64	9.92	12.94	3.72	8.70	6.73	5.48	3.11	100	

longest was 111 days. The length of stay was significantly different ($x^2 = 6235.319$, $P = 0.001$).

3.1.6. Treatment Outcome.
903 cases (5.67%) were cured clinically. 15011 cases (94.28%) improved; 8 cases (0.05%) were not cured. There was a significant difference in treatment outcome ($x^2 = 26688.02$, $P = 0.001$).

3.2. Analysis of Patients

3.2.1. Clinical Features.
There were 12,897 male patients in this group, accounting for 81.00%, which was 4.26 times that of female patients and was similar to that in Europe and America. The average age was (X34.97 ± 18.23) years old. The majority of patients were young adults aged 31 to 65 (49.21%), the oldest was 93 years old and the youngest was 5 days old. Preschool children aged 0–6 accounted for 12.27%. The injury time was the lowest in February (5.63%)

- thumb
- forefinger
- middle finger
- Ring finger
- little finger
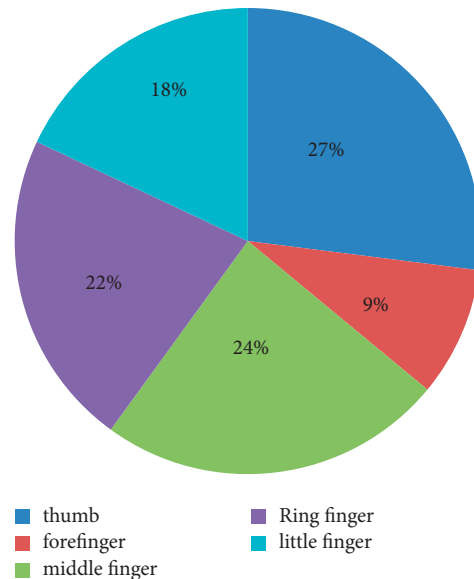
FIGURE 4: Distribution of finger injuries ($x2 = 785.87$, $P = 0.001$).

TABLE 3: Distribution of injury types.

Classify	Detached injury	Crushed injury	Concis	Frustration crack	Puncture injury	Fracture	Burn scald	Other	Total	Statistical analysis
Occupational	976	3924	3251	201	626	1021	198	669	10866	$t = 2.701$, $P = 0.031$
Living injury	401	1515	1436	162	705	501	157	179	5056	$t = 3.212$, $P = 0015$
Total	1377	5439	4687	363	1331	1522	355	848	15922	
(%)	8.65	34.16	29.44	2.28	8.36	9.56	2.23	5.33	100	
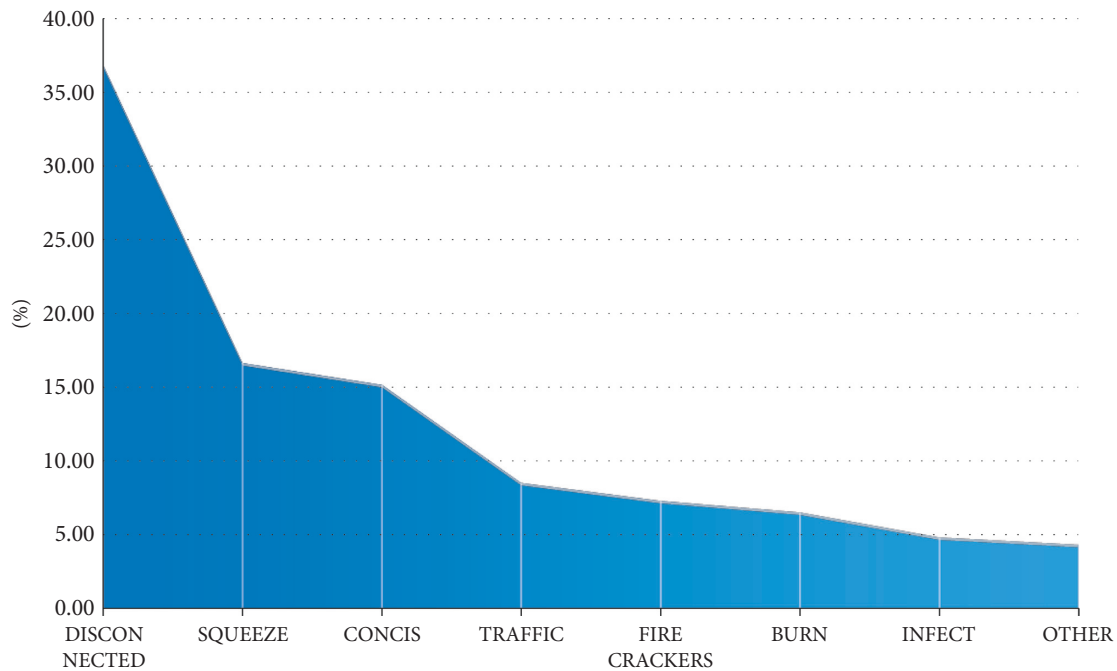
FIGURE 5: Causes of injuries of preschool children ($x2 = 1286.049$, $P = 0.001$).

and the highest in April (10.16%). The city's manufacturing industry is relatively developed, and the injured environment is mainly indoors such as factories, workshops, shopping malls, and service places (61.98%). The left hand was injured more than the right (1.16 : 1), and the left finger was injured more than the right finger (1.2 : 1). The thumb was injured the most (27%), followed by the middle finger (24%) and the ring finger (22%). The most common types of injuries were crush injury (34.16%) and cutting injury (29.44%). The average hospital stay of the patients in this group was 11.22 ± 12.75 days, and 79.41% of the patients were hospitalized for 1 to 2 weeks. The clinical cure rate was 5.67%, and the improvement rate was 94.28%. Eight patients were discharged from the hospital without healing due to concurrent infection, necrosis of the replanted finger, and nonunion of severed finger fractures.

3.2.2. *Occupational Injuries.* Hand trauma is a common surgical disease, and with the development of industrial machinery production, the incidence of acute occupational hand trauma is increasing [20]. Occupational injuries accounted for 68.25% of the cases in this group, of which the ratio of male to female was 5.23 : 1. There were more people aged 20–50 (74.18%), and men and young adults were the high-risk groups for occupational hand trauma. The main types of injuries were crush injury (36.11%), cutting injury (29.92%), and fracture (9.40%).

3.2.3. *Injuries of Preschool Children.* In this group, preschool children accounted for 12.27%, and the male: female ratio was 1.96 : 1, of which the 2–3-year-old group accounted for the highest proportion (45.70%). The main cause of injury was finger breakage (36.85%), followed by finger crush injury (16.63%) and cutting injury (15.15%).

3.3. *Precautions*

3.3.1. *Improve Manufacturing Intelligence.* In recent years, with the rapid development of IT in China and the rapid popularization of AI, occupational hand injuries have been gradually decreased. However, still the development is unbalanced. Some of the larger industrial and mining

enterprises have a low level of mechanical intelligence, insufficient protective facilities in the workplace, and the incidence of occupational injuries remains high. Therefore, it is recommended that the regulatory authorities improve the transformation and upgrading of equipment in large industrial and mining enterprises. At the same time, increase the promotion and application of AI technology, further improve the intelligence of the manufacturing industry, strictly enforce the access system, and reduce occupational injuries.

3.3.2. Improve the Comprehensive Quality of the Whole People. According to research in [21], there is a correlation between occupational hand trauma and risk factors such as contact with machine failure, change in operation method, sudden discomfort, lack of protective equipment, use of sharp knives, unfamiliar tools, unfamiliar machines, change of job, temporary overtime, temporary rush to work, and sudden physical discomfort during the dangerous period (30 minutes prior to the injury). The risk factors of occupational injury are closely related to many factors such as gender, age, occupation, cultural literacy, safety awareness, operational proficiency, working environment, conditions, and psychological status. It is recommended to strengthen vocational education, improve professional quality, strengthen safety awareness, improve the working environment, strengthen humanistic care, care about the physical and mental health of workers, and prevent safety accidents.

3.3.3. Strengthen the Supervision of Infants and Young Children. Traumatic injury of infants and young children has become one of the major problems in the social and public health fields of our country. To carry out in-depth clinical research on hand trauma in infants and young children, understand its clinical characteristics, strengthen preventive measures, reduce its incidence, and reduce the disability rate, should be raised as a systematic project to care for people's livelihood in today's society. This study showed that the proportion of hand trauma in preschool children was higher (12.27%), and the 2–3-year-old group was the main injury target of preschool children (45.70%). Because during this period, children appear as appearances, and it is also a period when their desire to imitate is high, showing the buds of exploration and knowledge, and they will make some inexplicable actions because of a strong need for independence, so it is easy to cause injury. Therefore, this stage of children is also an extraordinary period of parental supervision. In addition, the number of hand injuries in infants and young children is increasing year by year, which is also directly related to the increase of left-behind children in rural my country, or the fact that parents bring their children to work in cities, the ability to monitor infants and young children is insufficient, or they directly bring infants and young children into the work area [22]. Therefore, government care, social support, and the strengthening of guardianship are important measures to reduce accidental injury to infants and young children.

4. Conclusion

According to this study, males are more likely than females to sustain hand trauma; occupational injuries are more common than living injuries; males between the ages of 20 and 50 are at an increased risk of suffering an occupational injury. The incidence rate is highest in April every year and lowest in February. It mainly occurs in construction sites and factory workshops. The main types of injuries are crush injuries and cutting injuries. The main types of injuries were crush injury (36.11%), cutting injury (29.92%), and fracture (9.40%). Accidental injury of preschool children is also a social problem that cannot be ignored. This study showed that the proportion of hand trauma in preschool children was higher (12.27%), and the 2–3-year-old group was the main injury target of preschool children (45.70%). Therefore, measures such as improving the working environment, increasing the intelligence of the manufacturing industry, strengthening safety awareness, strengthening government functions, improving the quality of the whole people, strengthening humanistic care, and increasing infant care will greatly reduce the prevalence of hand trauma. To further extension of this research study, soft computing approaches can be adapted to predict the hand trauma injuries in the occupational environment [23, 24].

Data Availability

The data used to support the study are included in the paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] P. Angermann and M. Lohmann, "Injuries to the hand and wrist. A study of 50, 272 injuries," *Journal of Hand Surgery*, vol. 18, no. 5, pp. 642–644, 1993.

[2] S. Luria, D. Talmud, I. Volk, M. Liebergall, and R. Calderon-Margalit, "The epidemiology of wrist and hand injury in two hospitals in Jerusalem: substantial differences between population subgroups," *Israel Journal of Health Policy Research*, vol. 8, no. 1, p. 7, Article ID 6327559, 2019.

[3] W. H. Frazier, M. Miller, R. S. Fox, D. Brand, and F. Finseth, "Hand injuries:incidence and epidemiology in an emergency service," *Journal of the American College of Emergency Physicians*, vol. 7, no. 7, pp. 265–268, 1978.

[4] F. D. Burke, J. J. Dias, P. G. Lunn, and M. Bradley, "Providing care for hand disorders: trauma and elective," *Journal of Hand Surgery*, vol. 16, no. 1, pp. 13–18, 1991.

[5] M. S. Yardim, N. Cilingiroglu, and N. Yardim, "Financial protection in health in Turkey: the effects of the health transformation programme," *Health Policy and Planning, Health Policy and Planning*, vol. 29, no. 2, pp. 177–192, 2014.

[6] J. Urrutia, P. Besa, and F. Bengoa, "A prevalence study of thoracic scoliosis in Chilean patients aged 10–20 years using chest radiographs as a screening tool," *Journal of Pediatric Orthopaedics B*, vol. 27, no. 2, pp. 159–162, 2018.

[7] K. T. Yeh, R. P. Lee, I. H. Chen et al., "Are there age- and sex-related differences in spinal sagittal alignment and balance

among taiwanese asymptomatic adults?" *Clinical Orthopaedics and Related Research*, vol. 476, no. 5, pp. 1010–1017, 2018.

 [8] G. S. Sorock, D. A. Lombardi, R. B. Hauser, E. A. Eisen, R. F. Herrick, and M. A. Mittleman, "Acute traumatic occupational hand injuries: type, location, and severity," *Journal of Occupational and Environmental Medicine*, vol. 44, no. 4, pp. 345–351, 2002.

 [9] W. Kong and G. Xiao, "Current situation and prospect of epidemiological research on acute occupational hand injuries," *Journal of Environment and Occupational Medicine*, vol. 27, no. 8, pp. 501–504, 2010.

[10] D. Yuan, "Epidemiological study and prevention and control of hand trauma [J]," *Journal of Environment and Occupational Medicine*, vol. 27, no. 7, pp. 447-448, 2010.

[11] xiumei Jia and liqin Lan, "Investigation of 732 cases of occupational hand injuries," *Chinese Journal of Surgery of Integrated Traditional and Western Medicine*, vol. 21, no. 1, pp. 61–63, 2015.

[12] Z. Zhang and F. Wang, "Analysis of injury characteristics of occupational hand injuries in a hospital from 2013 to 2015," *Journal of Practical Preventive Medicine*, vol. 26, no. 3, pp. 327–329, 2019.

[13] N. M. Mule, D. D. Patil, and M. Kaur, "A comprehensive survey on investigation techniques of exhaled breath (EB) for diagnosis of diseases in human body," *Informatics in Medicine Unlocked*, vol. 26, Article ID 100715, 2021.

[14] sheng-zhi Zhang, Y.-jun Rui, and G. A. Smith, "Epidemiological analysis of emergency hand injuries in children in hospital," *Chinese journal of hand surgery*, vol. 33, no. 6, pp. 423-424, 2017.

[15] M. Kaur, S. R. Sakhare, K. Wanjale, and F. Akter, "Early stroke prediction methods for prevention of strokes," *Behavioural Neurology*, pp. 1–9, 2022.

[16] F. M0, B. C. K. Choi, C. Cilottey, L. Barbara, and R. Glenn, "Charecteristics and risk factors for accident injury in Canada from 1986 to 1996:an analysis of the Canadian Accident Injury Reporting and Evaluation (CAIRE)," *Injury Control and Safety Promotion*, vol. 9, no. 2, pp. 79–81, 2002.

[17] N. Chau, A. Bhattacherjee, and B. M. Kunar, "Relationship between job, lifestyle, age and occupational injuries," *Occupational Medicine*, vol. 59, no. 2, pp. 114–119, 2009.

[18] H. Grimsm0 Powney and E. C. Haris, "0ccupational health needs of commercial fishermon in southwest england," *Occupational Medicine*, vol. 60, no. 1, pp. 49–53, 2010.

[19] J. Urrutia, P. Besa, F. Narvaez, A. Meissner-Haecker, C. Rios, and C. Piza, "Mid and lower thoracic kyphosis changes during adulthood: the influence of age, sex and thoracic coronal curvature," *Archives of Orthopaedic and Traumatic Surgery*, vol. 142, no. 8, pp. 1731–1737, 2022.

[20] J. Naidoo, P. Govender, and D. Naidoo, "Taking hold of hand trauma in KwaZulu-Natal, South Africa," *African Health Sciences*, vol. 21, no. 4, pp. 1784–1793, 2021.

[21] W. Zhang and M. Kaur, "A novel QACS automatic extraction algorithm for extracting information in blockchain-based systems," *IETE Journal of Research*, pp. 1–13, 2022.

[22] T. C. Horton, J. J. Dias, and F. D. Burke, "Social deprivation and handinjury," *Journal of Hand Surgery*, vol. 32, no. 3, pp. 256–261, 2007.

[23] S. Wang, X. Liu, and H. Chen, "Case-crossover study on risk factors of occupational hand injury," *china occupational medicine*, vol. 41, no. 6, pp. 693–697, 2014.

[24] C. shouping, "epidemiological analysis of hand injuries in infants," *Chinese journal of hand surgery*, vol. 32, no. 2, pp. 101–103, 2016.

Research Article

Automatic Detection of Cases of COVID-19 Pneumonia from Chest X-ray Images and Deep Learning Approaches

Fahima Hajjej [ID],[1] **Sarra Ayouni,**[1] **Malek Hasan** [ID],[2,3] **and Tanvir Abir** [ID][4]

[1]*Department of Information Systems, College of Computer and Information Sciences,*
 Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[2]*Ajloun National University, Ajloun, Jordan, Iraq*
[3]*The University of Mashreq, Research Center, Baghdad, Iraq*
[4]*Department of Business Administration, Faculty of Business and Entrepreneurship, Daffodil International University,*
 Dhaka, Bangladesh

Correspondence should be addressed to Malek Hasan; malekb.has@yahoo.com

Machine learning has already been used as a resource for disease detection and health care as a complementary tool to help with various daily health challenges. The advancement of deep learning techniques and a large amount of data-enabled algorithms to outperform medical teams in certain imaging tasks, such as pneumonia detection, skin cancer classification, hemorrhage detection, and arrhythmia detection. Automated diagnostics, which are enabled by images extracted from patient examinations, allow for interesting experiments to be conducted. This research differs from the related studies that were investigated in the experiment. These works are capable of binary categorization into two categories. COVID-Net, for example, was able to identify a positive case of COVID-19 or a healthy person with 93.3% accuracy. Another example is CHeXNet, which has a 95% accuracy rate in detecting cases of pneumonia or a healthy state in a patient. Experiments revealed that the current study was more effective than the previous studies in detecting a greater number of categories and with a higher percentage of accuracy. The results obtained during the model's development were not only viable but also excellent, with an accuracy of nearly 96% when analyzing a chest X-ray with three possible diagnoses in the two experiments conducted.

1. Introduction

COVID-19, a new type of coronavirus, was discovered in lower airway samples from several patients in Wuhan, China, in December 2019 [1]. Fever, fatigue, a dry cough, and difficulty breathing were among the symptoms of severe pneumonia in these patients. It is believed to have originated in a seafood market in Wuhan. The spread of this virus created a situation that the World Health Organization classified as a pandemic (WHO). The incubation period was estimated to be 5.2 days, allowing the disease to spread globally via air travel. Evidence suggests that the virus can be transmitted during the incubation period in asymptomatic patients. [2] The virus can be spread from person to person through droplets from infected patients'

noses and mouths or through contact with contaminated surfaces.

More than 6 million confirmed cases of COVID-19 have been identified worldwide as of this writing, and multiple case reports have suggested that COVID-19 can be transmitted before symptoms appear. An infectious virus has also been found in presymptomatic COVID-19 cases [3]. While the level of infectious virus required for efficient transmission is unknown, detection of infectious virus extracted from upper airway samples suggests that COVID-19 transmission can occur before symptoms appear.

Many cities, state, and federal leaders have called for more real-time reverse transcription-polymerase chain reaction tests to check for the presence of RT-PCR (Reverse Transcription Polymerase) virus genetic material in response

to increased transmission. Researchers used chain reaction and serological assays to identify asymptomatic cases and potential spreaders [4].

Only people with severe disease symptoms are included in the Ministry of Health's current testing recommendations in Iraq. For the Iraqi population, the number of tests currently available is insufficient. Patients have been diagnosed after seven days, according to reports. The real-time Reverse Transcription Polymerase Chain Reaction (RT-PCR) test, which can detect SARS-CoV-2 RNA from airway samples, is the most common method for detecting COVID-19 cases. Although the RT-PCR test is highly specific, it is performed in a manual, time-consuming, laborious, and complicated process.

Furthermore, the current limitation of population-based tests in some countries, such as Iraq, emphasizes using auxiliary methods to identify COVID-19 positive cases. The X-ray examination, in which chest X-ray images (e.g., chest X-ray (CXR) or computed tomography (CT)) are taken and analyzed by radiologists to look for indicators, is another method of identifying COVID-19 [5]. SARS-CoV-2 virus infection causes visual symptoms. Patients with abnormalities on chest radiographs, typical of people infected with COVID-19, were found in early studies, suggesting that radiographic examination could be used as a primary tool for screening for COVID-19 in epidemic areas [6]. However, detecting pneumonia on chest radiographs is a difficult task for humans to complete; it is dependent on the availability of specialized radiologists, and it is done manually and takes time. The motivation for the work in question arises at this point, which aims to present a model for automatically detecting pneumonia due to COVID-19 from chest X-rays, which is a simple task for the model, which can detect with a high percentage of precision. It is a tool that can be used in conjunction with more conclusive tests as a supplement. Machine learning [7] is already being used as a resource for disease detection and health care as a supplement to help with a variety of problems that arise in daily life. The advancement of deep learning techniques and a large amount of data available enabled algorithms to outperform medical teams in certain imaging tasks, such as pneumonia detection, skin cancer classification, hemorrhage detection, and arrhythmia detection. The main concepts discussed in this work will be presented. Initially, the possibility of identifying cases of COVID-19 using X-ray images of the chest of patients will be described. Next, we talk about machine learning, deep learning, and convolutional neural networks, which are the basis of the study of the present work. Finally, the topic related to the VGG-19 architecture is addressed, which has its main applicability in the context of pattern recognition in images having a pretrained model.

1.1. Identification of COVID-19 from Radiography Images.
The majority of medical associations do not support the use of imaging modalities to screen patients with clinical suspicion of COVID-19. Computed tomography of the chest (CT) is recommended only for symptomatic hospitalised patients and portable chest radiography in particular

instances, such as inpatients who need imaging follow-up [8]. A normal chest CT scan does not rule out COVID-19, and neither does an abnormal exam confirm a clinical suspicion. An inexpensive, easy, and practical test for patients with suspected COVID-19 is a chest radiograph. Patients who are immobile or in specific conditions such as field hospitals can benefit from the method's portability, which can be used to monitor illness progression, evaluate tracheal tubes and drug infusion lines, and rule out problems like pneumothorax. Despite its availability and ease of performance, chest radiography (Figure 1) has low sensitivity in the evaluation of patients with clinical suspicion of COVID-19, ranging from 30 to 69%, with many tests being normal in mild forms of the disease. In altered exams, the main imaging findings are consolidations and low-density opacities, usually with peripheral basal predominance. Other findings, such as pleural effusion, are uncommon, and this finding was described in only 3% of patients in a study carried out by Wong et al. [9]. The peak of findings is seen around 10 to 12 days after the onset of symptoms, and the pulmonary changes may be of rapid progression, with the evolution to the middle and upper fields or diffuse pulmonary involvement, similar to that found in the diffuse alveolar damage of the syndrome of acute respiratory distress.

Patients with respiratory problems frequently have X-ray scans of their chests taken as part of their usual care. Because of the advantages listed above, portable chest radiography will become increasingly important as COVID-19 progresses.

1.2. Machine Learning.
Computers can learn without being explicitly programmed thanks to machine learning, which is a topic of research. Algorithmic induction is a step in the process of discovering new knowledge, and it is frequently referred to as "machine learning" when discussing this endeavour [10].

Induction is the process of creating a general model from a set of data. Induction can be associated with deduction; however, induction assumes a collection of facts and creates a general rule or model. Inductive learning is performed by reasoning about examples provided by a process external to the learning system.

The inductive learning process can be divided into the supervised learning process and the unsupervised learning process. In the supervised learning process, the inductor is provided with a set of training examples for which the associated class label is known. In general, the structure of each example is composed of a vector of characteristic values and the associated class label. The idea is to use this information in the induction algorithm to build a classifier that has the ability to identify the class of new unlabelled examples correctly. In the unsupervised learning process, the work works as follows: The inducer tries to determine if some of them can be grouped somehow from the given examples. After building these groups, an analysis is usually required to identify the context of each group within the situation being addressed.
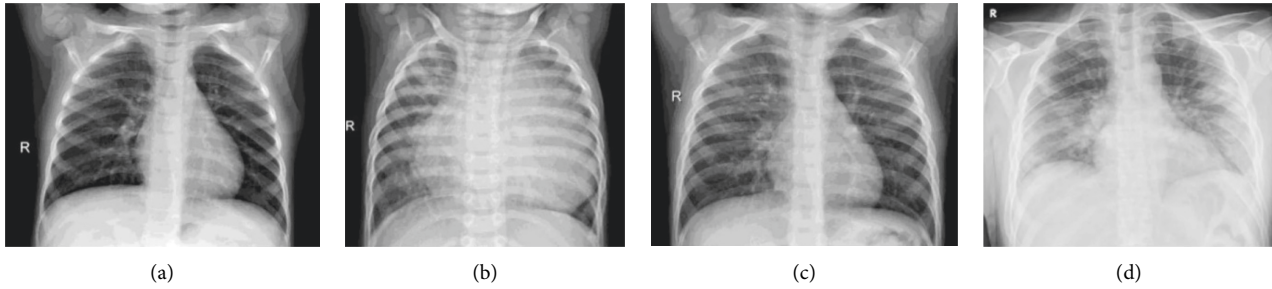
FIGURE 1: Example chest X-ray image of: (a) Healthy. (b) Bacteremia. (c) Viral pneumonia. (d) COVID-19 viral infection.
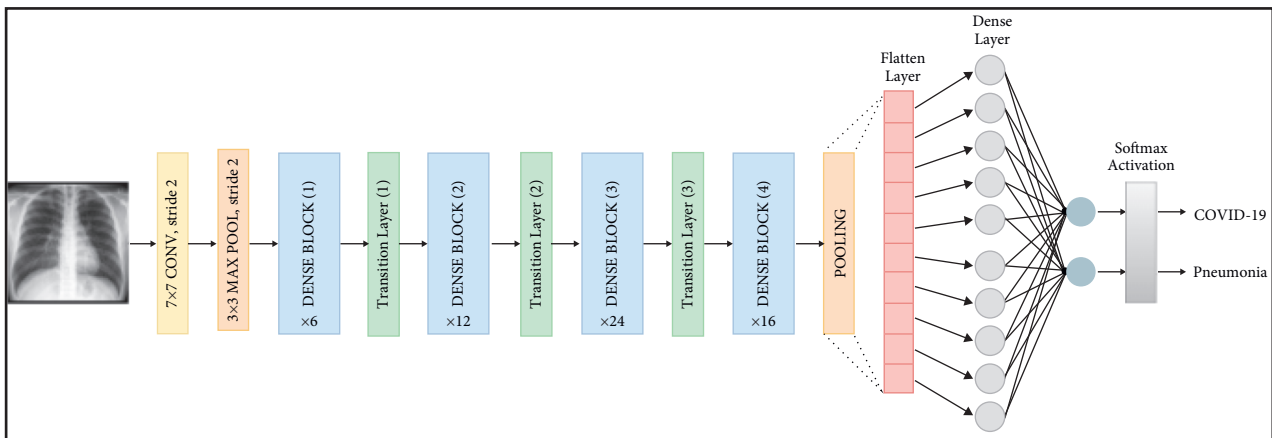
FIGURE 2: Building blocks of a traditional CNN.

Machine learning is at a very early stage in the process of discovering new information. Induction algorithms and other algorithms with the capacity to learn are the focus of machine learning, a scientific topic. In artificial intelligence, machine learning is a development in pattern recognition and computer learning theory. Human-like jobs can now be performed by machines, thanks to this research. But it is built on a foundation of previously learned principles that allow computers to make decisions based on a large amount of data.

1.3. Deep Learning. In machine learning, models are trained to perform useful tasks based on manually refined features from raw data. That is, it is necessary that the features used go through a feature engineering process so that it is possible to perform the classification or use features that were previously learned by another model. In deep learning, computers have the ability to identify useful features for the model automatically, directly from the raw data, bypassing the difficult step of manual information refinement. The main feature of the deep learning method is the focus on automatically learning data representations. This is the main difference between deep learning and traditional machine learning approaches. Resource discovery and task execution are merged into an issue and enhanced during the same training process [11].

1.4. Convolutional Neural Networks (CNNs). Deep learning has sparked attention as a result of the study and application of convolutional neural networks in medical imaging. The Convolutional Neural Network is a type of artificial neural network that preserves spatial correlations in data by having fewer connections across layers [12]. The convolutional network's layers keep track of the data's relationships as it is fed in. The architecture utilised in the technique without transfer of learning is more advanced than Duran-Lopez et al. model, but less advanced than Ozturk et al. model. The model utilised was proposed in Toni work and was chosen because it is a middle architecture between the two works, allowing for a better balance between overfitting and underfitting. Each layer operation operates on a small region of the previous layer as shown in Figure 2, the flow of a traditional convolutional network. These networks enable highly efficient input data and are very effective in image-oriented tasks. In the architecture of a CNN, several layers of activation and conversion operations are interleaved. Its training process is done using backpropagation and gradient descent.

1.5. VGG-19. The VGG architecture was proposed based on the AlexNet architecture–the architecture with the best performance in ILSVRC 2012 and has become a milestone in the use of Convolutional Neural Networks [13]. The main

VGG - 16
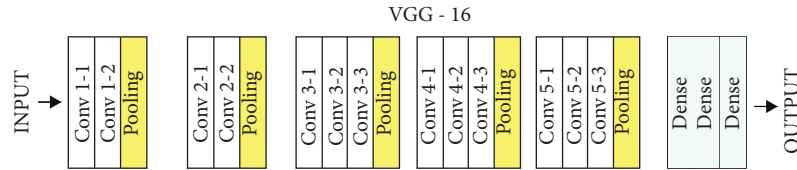
FIGURE 3: Example of the VGG-19 model network architecture.

feature covered in is depth. It is shown that a greater number of features can be extracted using architectures with more layers of convolutions and small convolutions (3×3, in this case). As a result, the VGG-16 architecture–16-layer VGG with synaptic weights–achieved great results at the 2014 ILSVRC, and although it did not win the competition, it has become one of the most used architectures in recent years, both VGG-16 and VGG-16, VGG-19. This architecture receives an RGB image of dimensions 224×224. ImageNet's 1.2 million general images with more than 1,000 different categories were merged into the dataset to train the VGG-19 model's 143 million parameters. The VGG-19 has a total of 19 layers for training, including convolutional and fully connected ones. The CNN VGG-19 architecture is shown in Figure 3.

2. Materials and Methods

This section presents the dataset chosen for the COVID-19 identification task and the proposed methodology. The model's training used chest X-ray images of patients with COVID-19, bacterial, and healthy pneumonia obtained from two distinct datasets available in the Kaggle repository: 'Chest X-ray Images (Pneumonia)' and 'COVID-19 chest X-ray'. The dataset was grouped into three categories: "COVID-19", "Bacterial", and "Healthy". In addition to identifying COVID-19, the model proposes to identify the presence of bacterial pneumonia. The method used proposes the use of the VGG-19 convolutional neural network, which uses as a basis for training the dataset created containing the three categories described above. One of the main objectives of the work is to identify a neural network architecture capable of identifying COVID-19 with a high percentage of accuracy. In addition, the work aims to fill a gap in the literature, which consists of identifying the presence of COVID-19, bacterial pneumonia, or a healthy state. During the creation of the model, two pieces of training were carried out using the dataset, but with different volumes, which will be described in this section.

2.1. Datasets. Two datasets were created to be used as a database set for training. The sets were created using two different chest X-ray images of patients. A given dataset x has images of viral pneumonia, bacterial, and healthy pneumonia, and dataset y have images of COVID-19. The x dataset, called Chest X-ray Images (Pneumonia) and the COVID-19 chest X-ray dataset, are available on Kaggle [14]. Kaggle is a community of data scientists with several datasets available for studies, such as those used in this work. Datasets created from these distinct sources have three

categories labelled "Covid-19", "Pneumonia-Bacterial", and "normal" in each case 266 images in dataset I and 288 images in data II set. What differs from Dataset I to Dataset II is just the volume of images contained in each one. During the selection of the images, tests were carried out to verify images that would reduce the performance.

During the selection of the images, tests were carried out in order to select images with higher quality to achieve better results during training. Images that reduced the model's performance were identified. Within the Dataset available in Kaggle, there are images with noise, noise that negatively influenced the accuracy during the tests. The selection was performed to discard images that were not frontal of the patients' chest and discard images with other noises, such as images that contained watermarks and digital arrows. Figure 4 exemplifies images that were discarded during the image selection process to compose the dataset.

2.2. Image Processing and Augmentation. The images that have been selected are loaded into a list in memory where they are converted to "RGB" format using the python pillow lib and resized to a new dimension of 150×150 pixels. The images are all labelled with their respective categories. Using a number of data augmentation techniques, the dataset was artificially extended and improved. The data augmentation strategy avoids overfitting and improves the model's capacity to generalise during training. Data augmentation is a data manipulation technique that generates new data instances without removing the essence or core of the original data. The settings used to increase the number of images are the Rotation range (20) and Zoom range (0.15) [15–20].

The rotation interval indicates the interval in which the images were randomly rotated during training, 20°. The zoom range randomly enlarges images by 15%. In the construction of this work, the treatment of image data with the *ImageDataGenerator* 3 object of Keras was used. *ImageDataGenerator* allows you to preprocess and augment the image dataset artificially in real-time during training. Only the data augmentation techniques provided by *ImageDataGenerator* were used in this work. This is very useful, especially when the dataset is very small. The rotation and zoom techniques used in this work effectively produce more data for training [20–22]. During training, all original images are transformed over the training periods with the configuration defined in the *ImageDataGenerator* creation parameters. In each epoch, the number of images is the same as that originally entered in the training dataset, but it undergoes real-time transformations during the training. When executing this procedure, new data are created artificially for training.
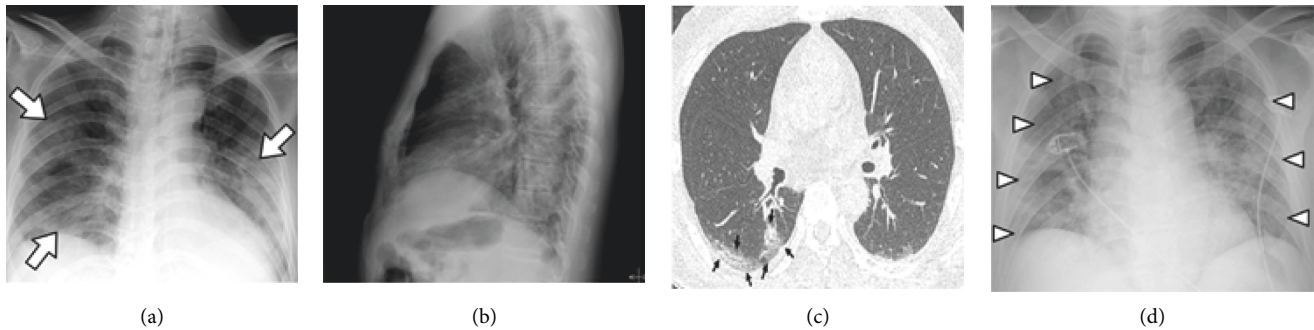
FIGURE 4: Example of images discarded in the selection process because they contain noise.

All experiments performed in this work were trained using 100 epochs as a parameter. It is possible to say that 100 different versions of each original image were used during the training. The created images are not completely different from the original ones, but they have a wide variety of modifications when rotated and zoomed in on. Because it is trained on a variety of versions of the same image, the resulting model will be more robust and accurate.

2.3. Segregation of Images. The segregation of images for training and testing is performed randomly, where 80% of the total dataset of chest X-ray images is destined for training the model, and the remaining 20% is used for its testing.

2.4. Neural Network Architecture. The model was created to identify the proposed problem having VGG-19 as a base, which is a convolutional neural network architecture with very small convolution filters (3×3) and with 19 layers in the base model, with the weights calibrated by the image net configuration.

2.4.1. Callback Functions. The ReduceLROnPlateau function was added to the model, which is a callback function that helps to reduce the learning rate by the factor, if there is no change in the loss. The configuration made for this function was to monitor the accuracy value given a configurable time X. In the experiment performed, this configured time was 2. If this time is reached without changing the accuracy value, the ReduceLROnPlateau will be applied given the configured factory. In the experiment, the factor configured for the function was 0.3.

2.4.2. Hyperparameters. During training, hyperparameters can be used to customise many components of the learning algorithm, which can have an impact on the final model's performance and accuracy. The hyperparameters used in the model tested in this work are batch size (34); Input Shape ($150 \times 150 \times 3$); Random State (42), and 100 epochs. The Batch size parameter configures the number of images per batch in processing. The value of the input shape parameter is used to configure the input shape. Given the configured value, the random state parameter is used to reproduce the experiment. Alpha is the parameter that determines the

TABLE 1: Compilation parameters.

Parameter	Configuration
Adam	Optimizer
Loss	Binary cross-entropy
Metrics	acc

learning rate and finally, the Epoch parameter configures the number of times the model will be trained.

2.4.3. Model Layers. This section describes the actual construction of the model, detailing the layers that were added. First, the sequential model was created by adding the VGG-19 model as a base.

The pooling configured for the model was made from a resolution by the global average of two dimensions. Adding batch normalization to the process increases the stability of the neural network by applying normalizations in the middle of training. The Flatten configuration was another layer added to the configuration, which serializes the image to the dense layer. A dense layer was also added to the construction of the model, adding 64 neurons with RELU activation. A dropout with a value of 0.4 was used in the configuration of the model; the dropout serves to improve the generality of the network. The last layer added was a dense layer with 3 neurons with softmax activation [23, 24].

2.4.4. Model Compilation. The parameters used when compiling the model are shown in Table 1.

Adam optimization is an optimization configuration for stochastic gradient descent based on the adaptive estimation of first- and second-order moments. The purpose of loss functions is to calculate the amount that a model should seek to minimize during training; the parameter setting chosen was binary cross-entropy which calculates the cross-entropy loss between true labels and predicted labels. The metric is a function used to judge the performance of your model and in this case, it is based on the accuracy ("acc") of the model that measures the frequency of predictions that match the labels.

3. Experiments and Evaluation of Results

In this section, the experiments performed are described and their results analyzed. The first experiment was performed

(a)   (b)
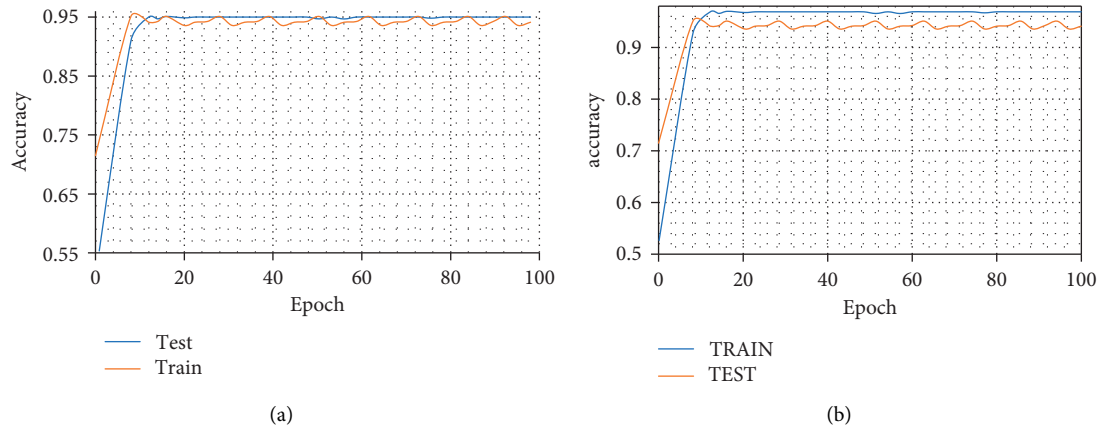
FIGURE 5: Model Accuracy. (a) First experiment. (b) Second experiment.

(a)   (b)
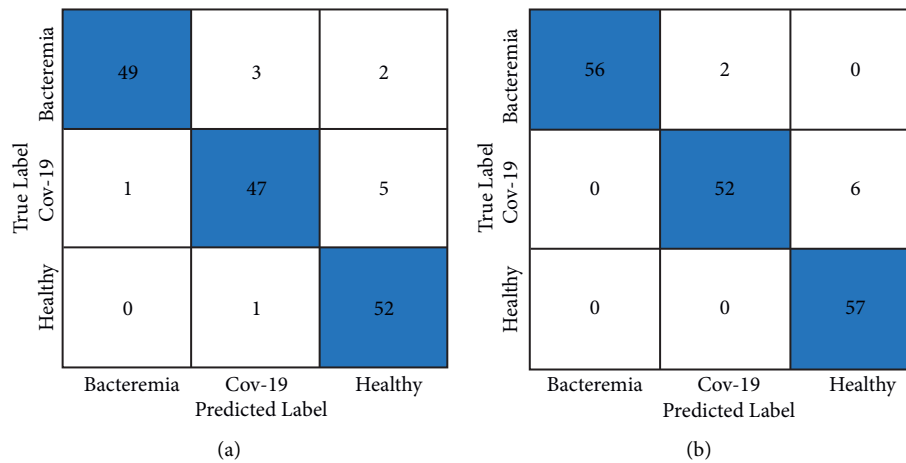
FIGURE 6: Confusion matrix. (a) First experiment. (b) Second experiment.

using "Dataset I" with 798 images for training. The second experiment was performed using the second training dataset, "Dataset II", with a total of 864 images. What differentiates the two experiments is only the amount of data offered for training the model. The metrics used to demonstrate the results of the two experiments were: Graph of the model's precision curve, which demonstrates the accuracy performance of the model training over the epochs; Confusion matrix, the confusion matrix demonstrates the complete classification accuracy of the model; Receiver Operation Characteristic Curve (ROC) and AUC Curve: For the classification case, the ROC curve demonstrates the model's performance in distinguishing positive and negative cases. The AUC curve is derived from the ROC curve ("area under the ROC curve"), and the result is arrived at by calculating the "area under the curve"; its value varies from 0.0 to 1.0. The higher the AUC, the better the model performance.

3.1. First Experiment.
In the task of automatically identifying COVID-19, Bacterial Pneumonia, or a healthy case through chest X-ray images of patients, the model proposed in his

first experiment used a dataset with a total of 798 images, 266 of which chest X-ray images of healthy patients, 266 images of COVID-19, and 266 images of bacterial pneumonia.

The data augmentation strategy was applied in real time and during the training, the model reached an accuracy of 0.9608. The accuracy model was also generated to demonstrate the performance during the training of the epochs Figure 5(a).

A confusion matrix was generated to demonstrate the test results Figures 6(a) and 6(b) in which the zero index means bacterial pneumonia; the index 1 means COVID-19 and the index 2 indicates a normal situation. Finally, to demonstrate the performance of the first experiment, the ROC curve metric is shown in Figures 7(a) and 7(b).

3.2. Second Experiment.
The second experiment used a dataset with 864 images for training, 288 images of chest X-rays of healthy patients, 288 images of COVID-19, and 288 images of bacterial pneumonia. The real-time data augmentation strategy was applied. During training, the model reached an accuracy of 0.9686. The accuracy model was generated to demonstrate the performance during the

(a)

(b)
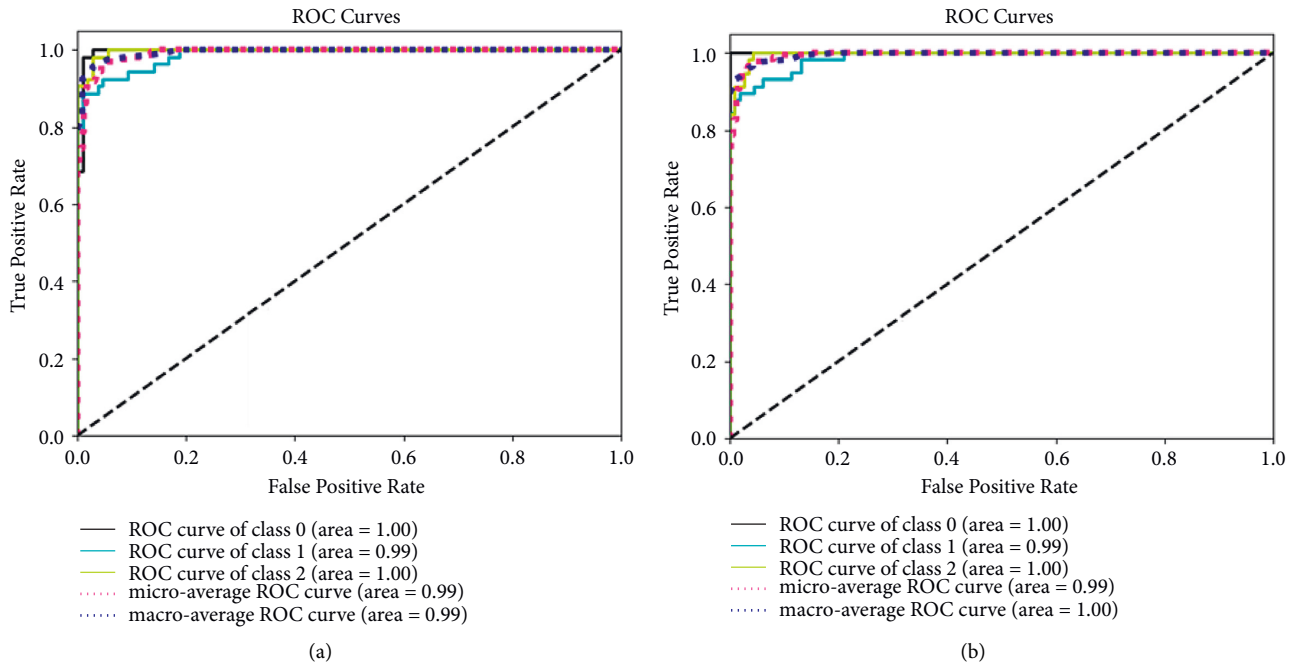
Figure 7: Receiver Operation Characteristic Curve (ROC). (a) First experiment. (b) Second experiment. (Class 0 = Bacteremia group; Class 1 = COVID-19 group and Class 2 = Healthy group).

training of the epochs Figure 5(b). To demonstrate the test results, a confusion matrix was generated.

(Figure 5) where index zero means bacterial pneumonia, index 1 means COVID-19 and 2 normal. Finally, to demonstrate the performance of the second experiment, the curve metric ROC Analysis of results is shown in Figure 7. In the task of automatically identifying COVID-19, Bacterial Pneumonia or a healthy case through X-ray images, a model was built with the configuration described in this work. The model can be subjected to training using datasets with different volumes as a data source. This was the focus assessed in the current study experiment.

The model was subjected to training using two datasets with different volumes. Dataset I was used with a lower volume than Dataset II to identify the model's ability to improve its accuracy when subjected to training with larger datasets.

The first experiment using Dataset I generated an accuracy result of 0.9608, in the second experiment using Dataset II presented an accuracy result of 0.9686. These results were created to provide a better view of the comparison of results between experiments.

4. Conclusions

Intending to develop alternatives for the diagnosis of COVID-19, which has proved to be a public health problem worldwide, and seeking approaches that made this diagnosis possible with the use of deep learning, the present work analyzed several articles as a way of evaluating the feasibility of developing the proposed project. After a thorough analysis, it is possible to state

that the imaging diagnosis of cases of severe acute respiratory syndrome based on X-ray exams is possible and has the possibility of differentiating it from bacterial pneumonia.

The results obtained during the development of the model proved to be viable and presented excellent results, with an accuracy of approximately 96% when analyzing chest X-rays with three possible diagnoses in the two experiments performed.

This work acts differently from the related works that were studied during the developed experiment. These works can perform a binary classification between two categories. An example is COVID-Net, which was able to identify a positive or healthy case of COVID-19 with an accuracy of 93.3%. Another example is CHeXNet, which can identify cases of pneumonia or a healthy situation in a patient with an accuracy of 95%. The current study, through experiments, proved to be efficient in detecting a greater variety of categories and with an even higher percentage of precision than the works cited. In future work, the objective is to carry out more experiments, more extensively validate the model developed and investigate the possibility of identifying other types of viral pneumonia to extend the diagnostic capacity of the model.

It will also be interesting to look for a dataset with a larger volume for training since the disease is recent, but a remarkable dataset is already available. However, with time, the tendency is to have more and more data to work within studies. Other artificial neural network architectures will also be implemented to evaluate possibilities for improvements, or even the development of a control system, to enable a safer diagnosis.

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.

Acknowledgments

References

[1] H. Zhu, L. Wei, and P. Niu, "The novel coronavirus outbreak in Wuhan, China," *Global Health Research and Policy*, vol. 5, no. 1, p. 6, 2020.

[2] W. Dhouib, J. Maatoug, I. Ayouni et al., "The incubation period during the pandemic of COVID-19: a systematic review and meta-analysis," *Systematic Reviews,*vol. 10, no. 1, p. 101, 2021.

[3] M. Mahmood, N. U. A. Ilyas, M. Khan, and M. Hasrat, "Transmission Frequency of Covid-19 through Pre-symptomatic and Asymptomatic Patients in AJK: A Report of 201 Cases," *Virology Journal*, vol. 18, no. 1, 2021.

[4] T. Suo, X. Liu, J. Feng et al., "ddPCR: a more accurate tool for SARS-CoV-2 detection in low viral load specimens," *Emerging Microbes & Infections*, vol. 9, pp. 1259–1268, 2020.

[5] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, "Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning," *Medical Image Analysis*, vol. 65, Article ID 101794, 2020.

[6] I. Y. A. A. A. R. Saad, "Social intelligence and its relationship to decision quality," *Scientific Journal Al-Imam University College, vol. 1, pp. 1-22, 2022. [7] M. F. Jwaid*, vol. 1, pp. 1–22, 2022.

[7] M. F. Jwaid and T. Baraskar, "An efficient technique for image forgery detection using local binary pattern (hessian and center symmetric) and transformation method," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–11, 2022.

[8] A. A. A. A. R. Baldawi, "The possibility of implementing industrial incubators and their role in the development of small industry and medium in Iraq," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–22, 2022.

[9] B. A. M. Muhammad, "The role of universities in developing societies by accreditation on scientific research," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–19, 2022.

[10] A. H. Abdullah, M. L. Thivagar, M. A. Bader et al., "Dynamic systems enhanced by electronic circuits on 7D," *Advances in Materials Science and Engineering*, vol. 2021, Article ID 8148772, 11 pages, 2021.

[11] W. H. A. Dahhan, M. H. A. Mashhadani, R. Raheem, E. Yousif, R. Raheem, and E. Yousif, "Iraq faces the COVID-19 with limited health capabilities and major medical challenges," *Bionatura*, vol. 5, no. 3, pp. 1271–1274, 2020.

[12] N. Ahmed, R. A. Michelin, W. Xue et al., "A survey of covid-19contact tracing apps," *IEEE Access*, vol. 8, pp. 134 577–134601, 2020.

[13] M. M. Rahaman, C. Li, Y. Yao et al., "Identification of COVID-19 samples from chest X-Ray images using deep learning: a comparison of transfer learning approaches," *Journal of X-Ray Science and Technology*, vol. 28, no. 5, pp. 821–839, 2020.

[14] W. Wang, X. Huang, Ji Li, P. Zhang, and X. Wang, "Detecting COVID-19 patients in X-ray images based on MAI-nets," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, p. 1607, 2021.

[15] R. M. Pereira, D. Bertolini, L. O. Teixeira, C. N. Silla, and Y. M. Costa, "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios," *Computer Methods and Programs in Biomedicine*, vol. 194, Article ID 105532, 2020.

[16] A. A. Hamad, M. L. Thivagar, M. B. Alazzam, F. Alassery, F. Hajjej, and A. A. Shihab, "Applying dynamic systems to social media by using controlling stability," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4569879, pp. 1–7, 2022.

[17] A. S. A. Obeidi and S. F. A. Azzawi, "A novel six-dimensional hyperchaotic system with self-excited attractors and its chaos synchronisation," *International Journal of Computing Science and Mathematics*, vol. 15, no. 1, pp. 72–84, 2022.

[18] H. M. Ahmed and B. W. Abdullah, "Overview of deep learning models for identification Covid-19," *Materials Today Proceedings*, vol. 15, 2021.

[19] J. V. Pranav, R. Anand, T. Shanthi, K. Manju, S. Veni, and S. Nagarjun, "Detection and identification of COVID -19 based on chest medical image by using convolutional neural networks," *International Journal of Intelligent Networks*, vol. 1, pp. 112–118, 2020.

[20] A. Karacı, "VGGCOV19-NET: automatic detection of COVID-19 cases from X-ray images using modified VGG19 CNN architecture and YOLO algorithm," *Neural Computing & Applications*, vol. 34, no. 10, pp. 8253–8274, 2022.

[21] E. Khan, M. Z. U. Rehman, F. Ahmed, F. A. Alfouzan, N. M. Alzahrani, and J. Ahmad, "Chest X-ray classification for the detection of COVID-19 using deep learning techniques," *Sensors*, vol. 22, no. 3, p. 1211, 2022.

[22] M. B. Alazzam, H. Mansour, F. Alassery, and A. Almulihi, "Machine learning implementation of a diabetic patient monitoring system using interactive E-app," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5759184, pp. 1–7, 2021.

[23] M. K. I. Rahmani, F. Taranum, R. Nikhat, M. R. Farooqi, and M. A. Khan, "Automatic real-time medical mask detection using deep learning to fight COVID-19," *Computer Systems Science and Engineering*, vol. 42, no. 3, pp. 1181–1198, 2022.

[24] G. Jain, D. Mittal, D. Thakur, and M. K. Mittal, "A deep learning approach to detect Covid-19 coronavirus with X-Ray images," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 4, pp. 1391–1405, 2020.

Research Article

Contradiction between Supply and Demand of Public Sports Services and Coping Strategies Based on the Genetic Algorithm

Liu Lu[1] and Wei Wei [ID] [2]

[1]*College of Physical Education, Chengdu Sport University, Chengdu 610041, China*
[2]*College of Sports Training, Chengdu Sport University, Chengdu 610041, China*

Correspondence should be addressed to Wei Wei; 20185658@m.scnu.edu.cn

The current situation of China's sports public service is not very satisfactory. This research analyzes not only the operation characteristics of the public sports service supply mode but also the connotation mechanism of the public sports service collaborative supply. The article also constructs the mechanism elements and the index system of the public sports service collaborative supply from the perspective of the genetic algorithm. Taking the eastern and central regions as research cases, this paper analyzes the actual situation of the collaborative supply of public sports services, and further explores the gap between urban and rural public sports. It puts forward the ideas and strategic paths for the innovative development of the collaborative supply model of public sports services. This research uses descriptive statistical analysis on the evaluation results of public sports service quality to examine the actual performance value of public sports service quality in various regions and uses the characteristics of different fitness of genetic algorithms to select the optimal supply and demand of public sports services in different regions. By analyzing the characteristics of insufficient supply and demand of public sports services in various regions, an optimal genetic algorithm model is constructed, and then combined with the scope of the study area and the number of people in the area, the genetic algorithm is used to continuously optimize until a set of optimal solutions appears, thus completing the optimal configuration of public sports services. The optimal path for the development of public sports in different regions is obtained, so as to promote the development of regional public sports. The study found that the satisfaction of sports funds in primary and middle schools in the east and middle is not very high, and the satisfaction with sports funds in middle schools is higher than that in primary schools, but only 50% of schools choose that sports funds are basically sufficient. The difference in satisfaction with sports funding in middle schools in the eastern and central regions is not obvious, and the satisfaction with sports funding in primary schools in the eastern region is higher than that in the central region. Judging from the amount of sports expenditures filled in by schools, 2000–3000 yuan per school year for primary schools and 4000–5000 yuan for middle schools are more common. Based on the calculation of 1,000 students in a school, the per capita sports expenditure is less than 10 yuan, which is very limited for sports training. This research will help to change the service concept and scientific decision-making and change the supply mode of sports public services. It will establish not only a vertical, complete, and horizontally smooth organizational structure and coordination mechanism but also a clear strategy for equalization of urban and rural sports public services.

1. Introduction

At present, the supply mode of urban public sports service in China is still in the initial stage of development. Therefore, many sports resources are difficult to play their potential role and cannot play their due role in improving the health and quality of life of the general public. The specific problems of

the supply of public sports facilities include insufficient supply of public sports facilities, single type, untimely community maintenance and management, low spatial combination, unbalanced and unreasonable layout, as well as insufficient number of social sports instructors. It is expected to promote the market-oriented operation process of China's public sports service supply, so as to further

optimize the allocation of urban public sports resources and realize the efficient and practical use of public sports resources. Under the background of service-oriented government, the focus of China's sports development is to promote the participation of all people in sports and improve the public sports service system. The purpose of this paper is to provide a basis for determining the supply goals and contents of public sports services in the future and to make certain references for the improvement of the national fitness program, sports functional departments, and social security service system.

Achieving the equalization of public sports services will help promote the development of sports in China. Xiangwei believed that the development of the city drives the progress of the entire economy and society. Meanwhile, the development of urban leisure sports has become an important engine of modernization [1]. Sun and Hu believed that people's average life expectancy has increased [2]. Chung aimed to elucidate how the organizational commitment and service orientation of sports practitioners are related to their job performance, with a view of helping sports practitioners provide good service and effectively organize their formation [3]. He believes that it is very important to promote comprehensive agricultural improvement [4]. Li believed that the services are a key link in the reform of the public service system [5]. Their proposed application of public sports services to the contradiction between supply and demand is not very effective for this situation. Through referring to the previous data, a genetic algorithm to optimize public sports services is proposed.

In the process of practice and application of the genetic algorithm, due to the limitation of population size and other factors, some excellent individuals will reproduce prematurely in the process of its evolution, thus reducing the diversity of the population. Tavakkoli-Moghaddam et al. proposed a GA algorithm [6]. Volkanovski A proposed a new method [7]. Long and Wu studied global optimization [8]. Gong et al. proposed an ensemble-based genetic algorithm [9]. Nemati et al. proposed an improved MILP method [10]. In the process of rapid industrialization, the agglomeration of population to cities has become the main source of "urban diseases," and community public sports services are facing severe challenges. Genetic algorithm has a better way of dealing with combinatorial optimization problems because it has better robustness, and the genetic algorithm will be further explored later.

The effective supply of public sports facilities is not only an important guarantee for the steady implementation of the national fitness strategy but also a powerful tool to solve the main social contradictions in the new era. This paper investigates the supply and demand of public sports services based on the genetic algorithm and finds out the problems existing in its supply based on the background of supply-side reform. The article not only analyzes the reasons for supply imbalance but also discusses the factors that lead to supply imbalance, as well as puts forward countermeasures and suggestions for supply-side structural reform. Starting from the purpose of building a service-oriented government, through interviews, comparative analysis, literature, field investigation, and mathematical statistics, the current situation of public sports service supply,

service consultation, facility construction, capital investment, activity development, and other issues and their effects are studied. Data collection and analysis are also carried out. The residents' public sports survey found that among the first, fitness paths accounted for 67%, far exceeding other sports venues. In the second place, outdoor fitness equipment and parks and squares accounted for 44%, which is the second main choice for residents to participate in sports activities.

2. Contradiction between Supply and Demand of Public Sports Services and Coping Strategies

2.1. Public Sports Services and Supply and Demand Equilibrium Theory

2.1.1. Public Sports Services. With the continuous development of society and the continuous growth of sports organizations, the role of sports activities in people's lives is becoming increasingly important, which plays a significant role in promoting people's health and improving people's quality of life. The government's single supply method is difficult to meet the needs of the general public in terms of public sports services, and it also makes many public sports resources idle and wasteful, making it difficult to give full play to the potential use value of public sports resources, which fails to achieve the original intention of serving the people. Under this situation, the government relaxes its powers and allows more social forces to enter the public sports service undertaking, which in turn produces a variety of public sports service supply methods. In particular, the market-oriented way is the most prominent, and it is also the most effective way of supplying public sports services. The market-oriented operation mode of China's public sports service supply is a reform of the traditional supply mode of China's public sports service. The reform is not simply the release and transfer of the government's power to the market, but the orientation of market-oriented reform as the core. With competition and monopoly, relying on market organizations and non-governmental organizations to achieve the diversification and efficiency of China's public sports service supply, only in this way it can move towards the marketization of public sports service supply [11].

The market-oriented operation of public sports service supply is the process of various activities to achieve the goal of public sports service supply. In this supply-demand relationship, the relevant government departments are the suppliers of it, and the masses are the demanders of it. The diversification of supply subjects enables the supply structure to be adjusted, which is also consistent with the supply-side structural reform. Meanwhile, as a supply product, public sports services are also diversified, such as public sports facilities, public sports information guidance, and public sports personnel guidance. Adjusting the structure of different types of public sports products also belongs to the category of supply-side structural reform. Strictly speaking, public sports service is the basic responsibility of the sports departments of governments at all levels, and it is the basis for implementing the outline of the national fitness plan and

carrying out various activities of national fitness. Only by continuously increasing the construction of public sports services can it have the conditions to truly implement the outline of the national fitness plan, so that increase in physical activity in China can be realized and the growing public sports needs of the masses can be met [12].

2.1.2. Supply and Demand Equilibrium Theory.
As a public sector, the government has the function of supplying those public goods that the market cannot provide or provide insufficiently, so as to achieve the purpose of making up for market defects. The characteristic of public goods is that every citizen of the society has the right to obtain it. The subject of its demand is the citizen, and its total demand is the sum of the demand of social citizens. When the public goods supplied by the government and the public goods demand of citizens are roughly in line with each other, the supply of public goods by the government and the public goods demand of citizens reach a state of structural equilibrium, but the structural equilibrium achieved is generally temporary and relative. In many cases, the public products provided by the government are not what citizens need, while the public products that citizens need are often not provided by the government; that is, supply and demand are in an unbalanced state. Even if it is an unbalanced state, it is limited by the scope of time and space, and it is an imbalance in line with the equilibrium theory of supply and demand [13].

2.2. Public Sports Service Supply System.
The composition of the public sports service supply system is shown in Figure 1. Providers generally refer to the fund providers of public sports products or public sports labor services, while producers generally refer to the specific executors of the production, processing, or public sports labor services of public sports products. When providing public sports services, there will be two situations in which the provider and the producer are the same or different. The situation where the provider is the same as the producer is called direct supply. Otherwise, it is called indirect supply.

Sports consumption classification is shown in Figure 2. In the classification of sports consumption, physical and non-physical sports consumption is divided according to whether the purchased product has a physical form. Spectator-type sports consumption refers to the behavior that citizens purchase sports tickets, sports performance tickets, etc., by paying money for the purpose of satisfying their own sensory stimulation needs, such as watching sports games or sports performances on the spot. Information-acquisition-based sports consumption refers to the behavior of citizens to purchase sports guidance services through payment currency to meet the needs of health and fitness, such as paying for fitness guidance or consultation.

2.3. Genetic Algorithms.
In the optimal allocation of the contradiction between supply and demand, the optimization process of the genetic algorithm is as follows: take the public sports service demand distributed to users in various regions as the decision variable, encode the decision variable to form the initial feasible solution set (i.e., generate several groups of initial

public sports allocation schemes), and then substitute it into the constructed optimization model to eliminate and select by judging the advantages and disadvantages of each solution, so as to generate a new generation of feasible solution set, repeat the optimization process until a set of optimal solution sets appear, so as to complete the optimal allocation of public sports services.

This study uses the fitness characteristics of genetic algorithms to improve and optimize public sports services. Genetic algorithm is a cyclic process, and new individuals are continuously generated through crossover and mutation of individuals. However, because the operation is random, the individual with the best fitness will be destroyed. The efficiency of the algorithm will be negatively affected. In order to keep the best individual, the optimal individual is stored in a space through the optimal individual preservation strategy, and it is used to replace the individual with the lowest fitness in the calculation. This strategy ensures that the optimal individual is not destroyed by genetic operations and enables the optimal individual to spread rapidly in the population. Exploring the ideas and strategic paths of innovative development of the collaborative supply model of public sports services through genetic algorithms.

The objective equation is as the following formula:

$$Q = \sum_{i=1}^{N} wt_i \sum_{k=1}^{K} wt_{i,k} \times Q_{i,k}^{j}, \tag{1}$$

where wt_i is the weight of the activity (i) in the project [14].

The crowding distance is as the following formula:

$$L[s]_{dis} = L[s]_{dis} + \frac{|L[s+1]m - L[s-1]m|}{L[k]m}. \tag{2}$$

In formula (2), $L[s-1]m$ is the mth objective function value of the sth individual.

Assuming that starting from $t = 0$, c remains constant, then formula (3) can be obtained as

$$m(H, t+1) = (1+c)^i m(H, 0). \tag{3}$$

The meaning of Cauchy distribution is that Cauchy distribution is a continuous probability distribution where mathematical expectation does not exist. When a random variable X satisfies its probability density function, X is said to obey the Cauchy distribution. The density function of the Cauchy distribution is as formulas (4) and (5):

$$f(x) = \frac{1}{\pi} \cdot \frac{t}{t^2 + x^2}, \quad -\infty < x < \infty, \tag{4}$$

$$F_i(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x}{t}\right). \tag{5}$$

For a given population of size n in generation t, the general description is as formulas (6) and (7):

$$f_{\max}(t) = \max\{f(a_1(t)), f(a_2(t)), \ldots, f(a_n(t))\}, \tag{6}$$

$$f_{\max}(t+1) = \max\{f(a_1(t+1)), f(a_2(t+1)), \ldots, f(a_n(t+1))\}. \tag{7}$$
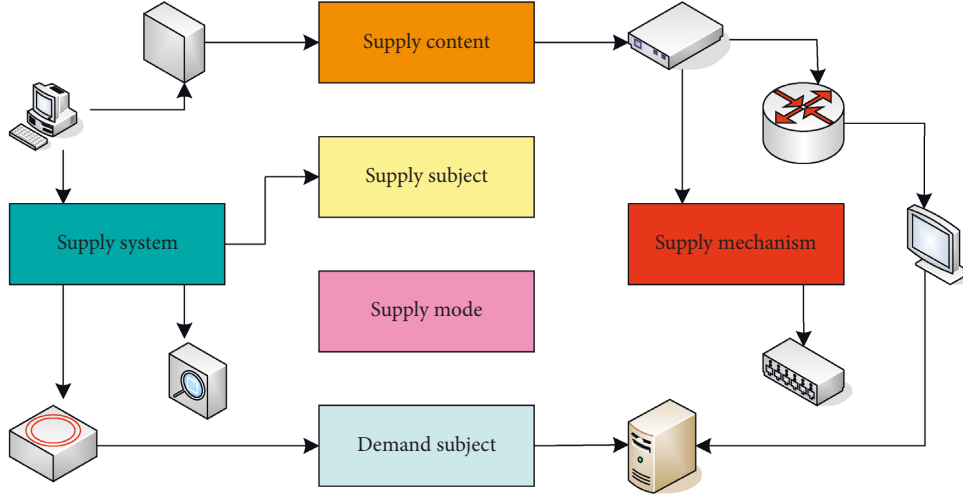
Figure 1: Composition of the public sports service supply system.
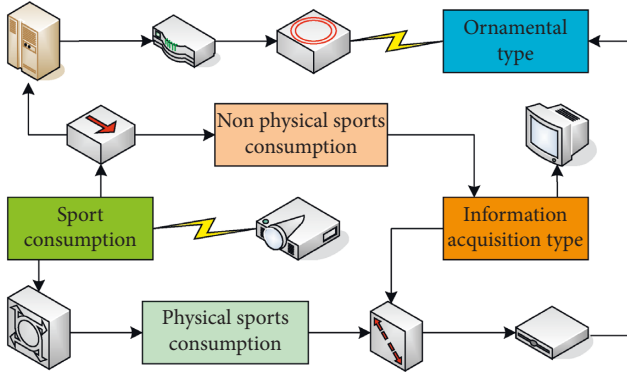
Figure 2: Sports consumption classification.

In dense blocks, the input of each layer is the output of the previous layer, namely,

$$x_n = H_n([x_0, x_1, \ldots, x_{n-1}]). \tag{8}$$

Given a sample set that satisfies the independent and identical distribution, the task of the restricted Boltzmann machine is to find the value of the parameter $\theta = \{w, a, b\}$, and then fit the given learning sample. Formula (9) can get through derivation of the maximum log-likelihood function [15]:

$$L(\theta) = \frac{1}{N} \sum_{n=1}^{N} Log p_\theta(v^n) - \frac{\lambda}{N} \|W\|^2. \tag{9}$$

When L is the largest, the corresponding parameter W is as the following formula:

$$\frac{\partial L(\theta)}{\partial W_{i,j}} = Ep_{\text{data}}[v_i h_j] - Ep_\theta[v_i h_j] - \frac{2\lambda}{N} W_{ij}. \tag{10}$$

Assuming that L represents the set of load balancing indicators, and the set $L = \{L_1, L_2, \ldots L_n\}$, then $Lj(X)$ represents the load indicator of cluster node j under the scheduling policy X. Then $Lj(X)$ defines as the following formula:

$$L_j(X) = \frac{1}{T(X)} \sum_{i=1}^{m} M_{ij} * W_{ij}. \tag{11}$$

Assuming that $\rho(A)$ represents the error of the load index, then $\rho(X)$ can be defined as the following formula:

$$\rho(X) = \sqrt{\frac{1}{n}\left(L_j(X) - \frac{1}{n}\sum_{j=1}^{n} L_j(X)\right)^2}. \tag{12}$$

At time t, the pheromone concentration of server node j is initialized as the following formula:

$$\tau_{ij}(t) = \frac{1}{LF_j}. \tag{13}$$

The pheromone (Pheromone, also known as phero-mone, refers to the secretion of an individual into the body, and is detected by other individuals of the same species through the olfactory organs (such as the accessory olfactory bulb and the vomeronasal organ). It is the substance that causes the latter to exhibit some behavioral, emotional, psychological, or physiological change. Concentration changes simultaneously with the heuristic information. At time t, the heuristic information is initialized as the following formula:

$$n_{ij}(t) = \tau_{ij}(t). \tag{14}$$

Pheromone update is defined as the following formula:

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \Delta\tau_{ij}(t, t+1). \tag{15}$$

In formula (15), ρ represents the pheromone volatilization coefficient.

At time $t+1$, the heuristic information changes as the following formula:

$$n_{ij}(t+1) = \tau_{ij}(t+1). \tag{16}$$

The execution time of each resource needs to be calculated to execute the task using a matrix such as decoding sequence. Therefore, the total time to complete the resource scheduling task is as the following formula:

$$F(x) = \max_{r=1}^{n} \sum_{i=1}^{W} \text{work}(r, i). \tag{17}$$

In formula (17), work (r, i) represents the time spent by resource r to execute subtask i on the resource, and W represents the number of subtasks assigned to the resource.

3. Research Objects and Method

3.1. Research Objects.
The selection of the sample adopts the method of random classification sampling. The subordinate communities of each sub-district office are selected separately, and the specific selection number is 16 community managers and 320 community residents.

3.2. Interview Method.
Interviews with relevant leaders of the National Sports Federation on the nature, management, operation, and funding sources of individual sports associations are conducted, and interviews with the relevant leaders of the Social Sports Section of the Sports Bureau on the construction of venues and equipment, management mechanisms, social sports instructors, and funding for community sports are also conducted. The development of community sports can be grasped and relevant information can be obtained. After communicating with the person in charge of the advanced sports unit on the organization and management of sports activities in the community, the source of funds, the participation of residents, the status of venue equipment, etc., first-hand information can be obtained. The research mainly adopts the combination of expert opinions and boundary delineation for index selection. The delineation of the threshold is to calculate the arithmetic mean (the arithmetic mean is the most basic and commonly used average index in statistics, and it is divided into simple arithmetic mean and weighted arithmetic mean), the frequency of full marks, and the coefficient of variation according to the selected scores of each indicator by experts. Each indicator has these three different discriminant scales.

The arithmetic mean calculation formula is as the following formula:

$$M_j = \frac{1}{n} \sum_{i=1}^{n} c_{ij}. \tag{18}$$

In formula (18), M_j refers to the arithmetic mean; n is the number of experts, and C_{ij} refers to the total number of scores of the ith expert on the j index.

The formula for calculating the full frequency is as the following formula:

$$K_j = \frac{m_j}{n_j}. \tag{19}$$

In formula (19), K_j is the full frequency, and the value of K_j is between 0 and 1.

The formula for calculating the coefficient of variation is as the following formula:

$$V_j = \frac{P_j}{M_j}. \tag{20}$$

3.3. Questionnaire Survey Method.
A questionnaire was made for the basic situation of community residents, willingness to participate in physical exercise, consumption level of sports, exercise venues, exercise items, sports organization management, and fitness guidance in their own communities. A total of 500 questionnaires were distributed to the residents of 3 streets in the central district of the city. Starting from the needs of this research, after conducting interviews with relevant experts, two questionnaires were designed. The design of the questionnaire strictly followed the selection of basic principles and methods. Relevant survey questions for the current situation of sports public services were preliminarily designed. After asking the relevant experts to revise and improve the content of the questionnaire, the relevant questionnaires were distributed, and the distribution objects were the actual managers of community sports and the two groups of residents participating in it. The purpose of the questionnaire design research was to understand the development status of the validity test of the questionnaire: in order to ensure the scientificity and validity of the stable questionnaire survey in this study, the designed questionnaire was sent to 10 experts for expert logic test. It was necessary to test the validity of the questionnaire including content, structure, and overall aspects. The investigation materials were classified and coded, and Spss21 is used to analyze data. SPSS21 Chinese version is a very powerful professional data analysis tool. SPSS21 version is convenient and easy to use. It adopts the input method and the management module similar to the Excel table, and data statistics is also extremely simple and convenient.

4. Simulation Experiment Analysis on the Contradiction between Supply and Demand of Public Sports Services

In this study, the actual valid questionnaire data were collated, and the results are shown in Table 1.

The system of inter-class sports activities needs to be fully implemented. 25–30 minutes of large-scale inter-class sports activities should be arranged every morning, and students should be seriously organized to do radio gymnastics. The collective sports activities needs to be carried out: boarding schools should insist on doing morning exercises every day. Table 2 shows the standards for sports venues in national and local compulsory education schools.

A total of 25 primary schools and 25 secondary schools in four eastern and central provinces were investigated for school physical education in this study. The study period for primary and junior high schools in China's compulsory

TABLE 1: Arrangement of the actual valid questionnaire data.

Gender	Number of people	Percentage (%)
Male	250	55.6
Female	200	44.4
Total	450	100

TABLE 2: National and local compulsory education school sports venue standards.

Area	Types of school	Number of classes	Track field
National standard	Primary school	≤18	200 meters (ring) 1 piece
	Junior high school	≤18	300 meters (ring) 1 piece
National rural standard	Primary school	≤6	300 m–400 m (ring) 1 piece
	Junior high school	≤18	300 meters (ring) 1 piece

education stage is the 6–3 system, the 5–4 system, the 5–3 system, and the nine-year consistent system, respectively. This article still uses the name, junior high school, to refer to the original junior high school grades 7, 8, and 9 of the nine-year system. When the name of middle school is used in the following text, it refers to junior high school unless otherwise specified. The situation of schools in different regions is shown in Figure 3.

Based on the "Construction Standard of Rural Ordinary Primary and Secondary Schools," which is a lower standard for track field materials, and the standard that every 6 classes should have 1 basketball court or volleyball court, the current situation of school sports venues in various regions of China was investigated (the sports venues and equipment in the eastern region are shown in Figure 4(a)). In the eastern and central regions of this survey, the compliance rate of the middle school track and field is higher than that of primary schools, and secondary schools that fail to meet the standards are usually due to insufficient field area. Primary schools that do not meet the standards are mostly because there is no track and field at all. The central region is mainly underconstructed (the sports venues and equipment in the eastern region are shown in Figure 4(b)). Judging from the materials of the track and field, the plastic track in the eastern region has become more popular, while the central region is still dominated by the cinder track. In addition, 8 primary and secondary schools in the eastern region have indoor gymnasiums, 5 in the central region, and 4 schools in the eastern and central regions have swimming pools. Judging from the situation of school basketball volleyball court, the standard rate of middle school basketball volleyball court is also better than that of primary school, and the middle and primary schools are better than the eastern middle and primary schools. This is mainly because schools with insufficient track and field construction usually use simple basketball and volleyball courts as venues for school physical education and daily teaching activities. In addition to basketball and volleyball courts, the most common sports facilities in most schools are ping pong tables. Except for schools with extremely poor sports facilities, almost every school has a certain number of outdoor table tennis tables.

The satisfaction with sports venues in middle schools is significantly higher than that in primary schools. The equipment satisfaction of the middle and primary schools in the eastern and central areas is similar, but the equipment satisfaction is slightly higher than the venue satisfaction (the sports equipment satisfaction survey in the eastern region is shown in Figure 5(a)). Sports equipment rooms are available in most schools. Although more than 70% of the schools in the equipment self-assessment believe that they have equipped sports equipment in accordance with the "Basic Standard for Trial Implementation of National School Sports Hygiene Conditions," the proportion is obviously high from the survey and visit of some schools. The proportion of schools that can be fully configured according to this standard is not too high. Most schools can configure some frequently used sports equipment that is not easily damaged, and the easily worn sports equipment is rarely updated even if it has been configured (the satisfaction survey of sports equipment in the central region is shown in Figure 5(b)).

The class-teacher ratio of primary and secondary schools is based on the standard of 6 classes with a full-time physical education teacher. Overall, only slightly more than half of the schools equipped with physical education teachers meet the standard. The allocation of PE teachers in middle schools is better than that in primary schools (The allocation of PE teachers in the east is shown in Figure 6(a)). Through the understanding of some eastern schools, some schools are more inclined to recruit main subject teachers although there is a shortage of physical education teachers. They usually let other teachers come to guest appearances in the case of insufficient physical education teachers. Therefore, the eastern school's emphasis on school physical education is not proportional to the development of education (the central PE teachers are equipped as shown in Figure 6(b)).

Taking 4 lessons per week for grades 1–2 in primary schools, 3 lessons per week for grades 3–6 and junior high schools, and 2 lessons per week for high schools, the survey found that the eastern region's physical education curriculum compliance rate was significantly higher than that in the central region. The situation of non-standard physical education in primary schools in the central region is relatively serious, especially the lack of physical education courses for grades 1–2 in primary schools. The arrangement of school morning exercise is related to the school accommodation system. Usually, the school with the
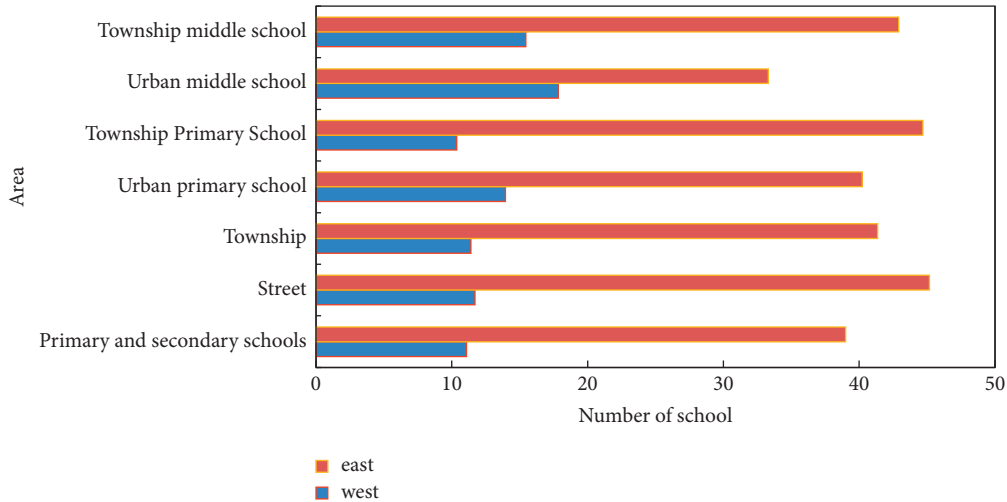
Figure 3: School situation in different regions.
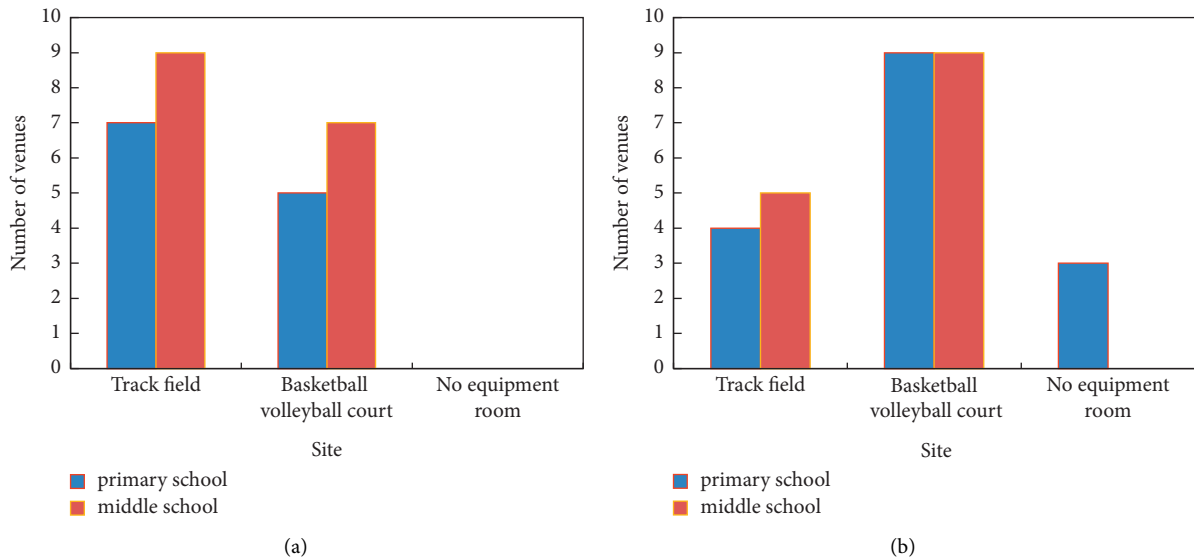
Figure 4: Sports venues and equipment in eastern and central regions. (a) Eastern region. (b) Central region.

accommodation system will carry out morning exercise activities, although during the implementation process, the morning exercise time becomes a signal for students to get up and prepare time before class. In the survey, most of the middle and primary schools in the Middle East could carry out and implement large-scale inter-class activities. From the perspective of the content of activities, they were mainly broadcast gymnastics (winter long-distance running). Some schools could carry out certain forms of sports activities, especially in the ninth grade. Because of the additional physical examination in the senior high school entrance examination, they paid more attention to quality practice during the large class. However, there were also some schools that did not implement large inter-class activities and only replace them with eye exercises. The extent of the development of large recesses is largely consistent with the conditions of school sports venues. The better the school venues and facilities are, the better the development of large recesses will be. Judging

from the situation of extracurricular activities arranged by schools, most of the middle and primary schools in the east and middle could carry out extracurricular activities. In terms of self-assessment of one-hour daily exercise time in schools, exercise varies from region to region, this difference was more due to the difference in physical education. According to the survey, primary school students were better than middle school students in exercising for one hour a day. The main reason for this situation is the continuous increase of pressure for further education. The eastern primary and secondary schools could generally meet the standard of one physical fitness test per school year, and the physical fitness test in middle schools was also better than that in primary schools. Due to the incomplete data of each school on the pass rate, obesity rate, and myopia rate of middle school students who returned the questionnaire, it was impossible to analyze the students' physical condition through this survey. From the limited data, secondary school myopia and the primary school obesity rate were higher. The
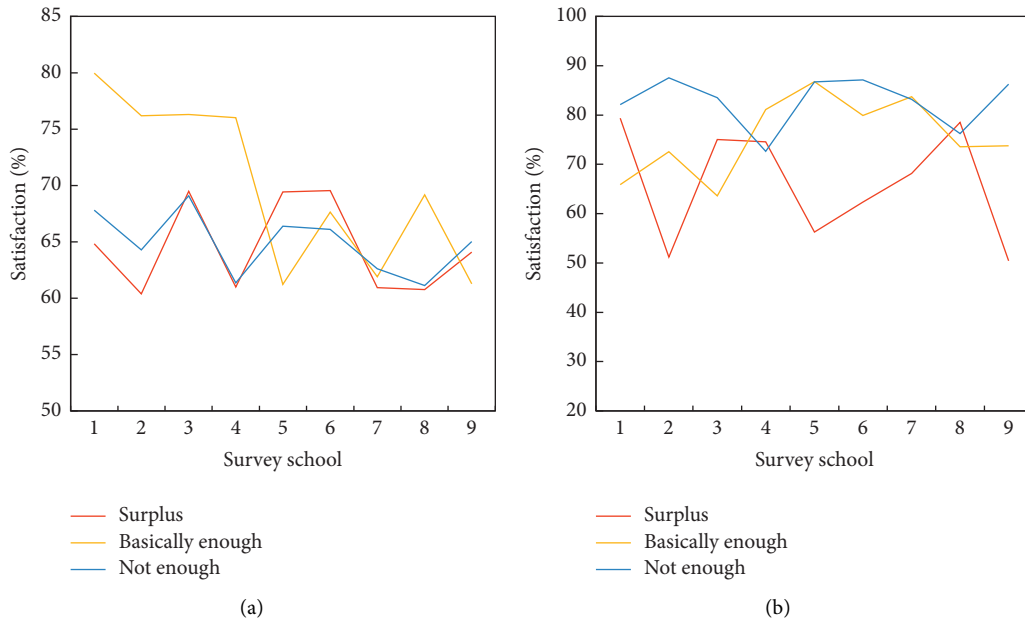
FIGURE 5: Sports equipment satisfaction. (a) Eastern region. (b) Central region.
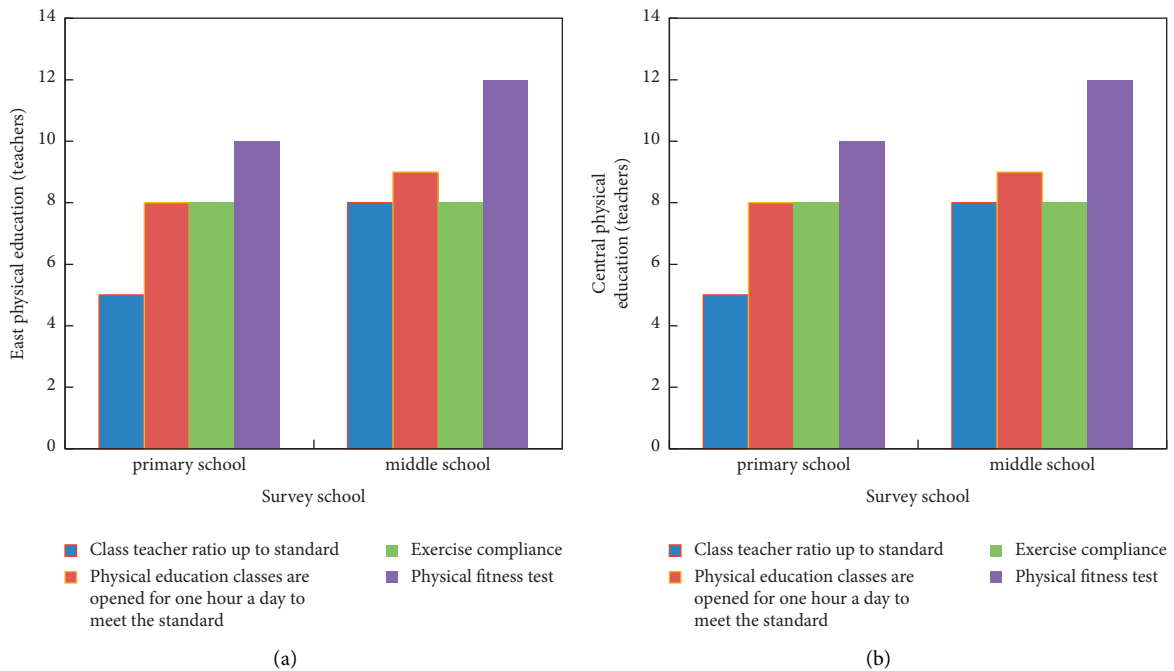
FIGURE 6: Physical education in primary and secondary schools. (a) Eastern PE teachers. (b) Central PE teachers.

middle and primary schools in the east and middle were generally able to hold a track and field meeting every school year, and a relatively high proportion could hold 1–2 intramural ball sports competitions and 3–4 other sports competitions per school year. The ball games are mainly basketball, volleyball, and table tennis, and other sports competitions are mostly tug-of-war and rope skipping.

Several other important issues in school sports work include the budget expenditure of school sports funds, the acceptance of sports support from government departments,

and the opening of school sports facilities. At present, China's primary and secondary education funds are mainly derived from public funds allocated according to the number of students in the school. It can be seen that the satisfaction degree of sports funds in primary and middle schools in the eastern and central regions is not high, and the satisfaction with sports funds in middle schools is higher than that in primary schools. Only 50% of schools choose sports funds that are basically sufficient. There is little difference in the satisfaction level of sports funding in middle schools in the
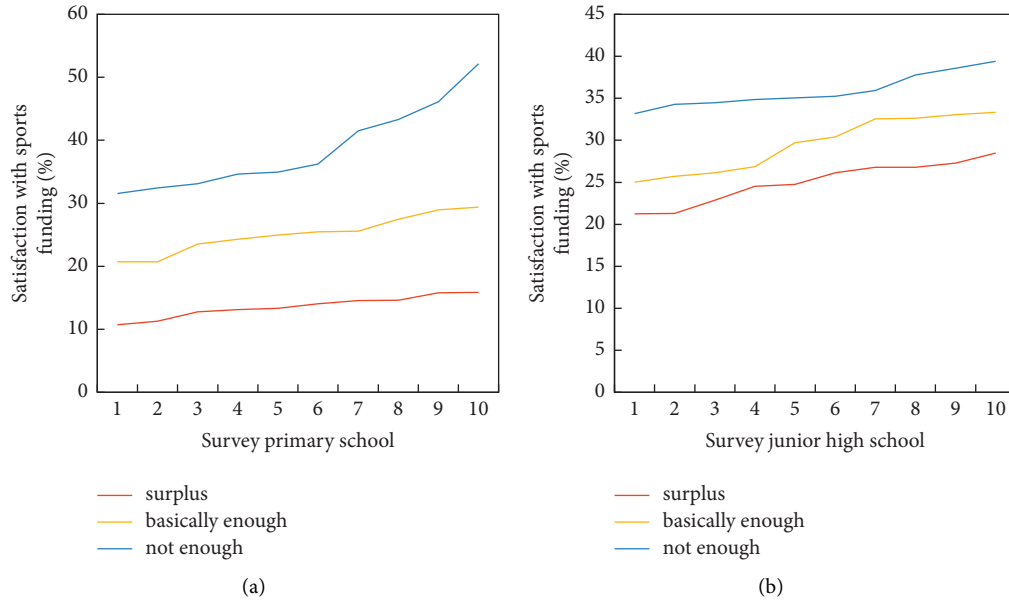
FIGURE 7: Results of the self-assessment survey on sports expenditures in primary and secondary schools in the eastern and central regions. (a) Elementary school. (b) Junior high school.

eastern and central regions. The satisfaction with sports funding in primary schools in the eastern region is higher than that in the central region (primary school funding is shown in Figure 7(a)). Judging from the amount of sports expenditures filled in by some schools, 2000–3000 yuan per school year for primary schools and 4000–5000 yuan for middle schools are more common. Even if a school has 1,000 students, the per capita sports expenditure is less than 10 yuan. Some schools with sports teams have spent 10,000–20,000 yuan in sports funds. Even if it is all used for daily sports work, the funds are quite limited, not to mention training and competition. Education, sports, and other government departments sometimes provide certain sports services and support to primary and secondary schools, mainly in the form of sports equipment donations and business training, but the proportion and frequency are not high (The funding for middle schools is shown in Figure 7(b)).

The issue of the opening of school sports venues and facilities to the public has been widely concerned. Judging from the results of this survey, except for some schools in the middle and primary schools in the east and the middle, they do not have venues and facilities and are unable to open. Some schools are open because the sports facilities and teaching areas can be separated in the area, and some schools and public sports facilities are actually open. The most that other schools can do is to open them to their own school and surrounding students for free after school and during holidays. Under the encouragement of the local government's school opening policy and financial subsidies, some schools open some indoor venues for a fee, but the overall utilization rate is not high because the price is significantly higher than that of socially operated sports venues.

The selection of residents' sports activities and types of facilities is shown in Table 3. Residents ranked the top three purposes of participating in physical activity according to their importance. Maintaining good health is the first purpose of

TABLE 3: Selection of resident sports activities and types of facilities.

Purpose of physical activity	1nd (%)	2nd (%)	3rd (%)
Stay healthy	64	26	4
Entertainment	23	26	46
Relax	31	36	31
Improve physical fitness	14	39	44

residents participating in sports activities, which is far ahead of other purposes, reaching 64%. Maintaining and improving the body is the second purpose of residents participating in sports activities, accounting for 44%, which is 6 percentage points higher than the relatively low goals of improving physical fitness and making friends at the same level. Recreation is the third purpose of residents participating in sports activities, accounting for 46%. However, from the further analysis of the average ranking, maintaining good health, improving physical fitness, and entertainment accounted for a relatively high proportion, ranking the top three. Except for the second main purpose, the other purposes are the same, indicating that the main purpose of residents participating in sports activities is consistent with the essential function of sports. Combining all the purposes of activities, it is found that the residents' sports interest has gradually shifted from direct interest to indirect interest, and the purpose of activities has diversified, which poses a challenge to the supply of community public sports facilities.

With the popularization of the concept of national fitness, people have gradually realized the role of physical fitness in life, but at the same time, the quality cannot be guaranteed from the perspective of supply side structural reform; this paper investigates the supply and demand of public sports services from the current situation, deeply analyzes the supply and demand contradiction of public sports services, defines the current supply and demand situation of public sports in China, uses the genetic algorithm to find out the contradiction between them,

TABLE 4: Selection of public sports facilities.

Fitness venue	1nd (%)	2nd (%)	3rd (%)
Outdoor fitness equipment	11	44	42
Indoor gym	40	35	57
Outdoor fitness venues such as parks and squares	20	67	27
From home yard	27	23	45

and puts forward the solution of marketization, socialization, organization, and legalization of public sports services.

Due to the different sports preferences of residents and the range of activities, the selection of public sports facilities is shown in Table 4. Through sorting and analyzing the top three important places where residents often participate in sports activities, it is helpful to determine the characteristics of residents participating in sports activities. In the first place, fitness paths accounted for 67%, far exceeding other sports venues. In the second place, outdoor fitness equipment and parks and squares accounted for 44%. In the third place, other types of sports venues accounted for 57%, which was mainly due to the diversity of residents' sports activities determined by the diversity of residents' sports venues. Secondly, the average ranking analysis of residents' top three sports venues show that outdoor fitness equipment, other sports venues, and chess and card venues are the top three residents' sports venues on average. It can be seen that there is a large discrepancy between the importance ranking and the average ranking, and the importance ranking and the average ranking of other sports venues are higher, indicating that community residents participate in various types of sports activities. However, the overall analysis of free venues provided by the government is still the main choice for residents to engage in sports activities. This has a certain relationship with the income of residents. Genetic algorithm can reasonably allocate the optimal activity type for residents according to the characteristics of residents' sports activities, the stadium, and the number of activities.

Judging from the survey, the urban primary and secondary school track and field and basketball and volleyball courts are slightly better than the towns. The reasons for the great difference between school stadiums and gymnasiums are as follows: on the one hand, the streets are located in the urban area, there is no guarantee of constructing land for sports venues, so the sports venues cannot meet the standards. On the other hand, township schools, especially those where the local government is located, usually do not lag behind in the construction of school infrastructure, but the school sports facilities in rural areas outside the location of the township government are not guaranteed at all. Meanwhile, the vast majority of plastic track and field venues are located in urban primary and secondary schools. As far as the configuration of outdoor table tennis tables is concerned, there are slightly more schools in townships than schools in urban areas, and the shortage of land for schools in urban areas is one of the reasons. Another reason is that both urban schools and students have higher requirements for sports facilities than townships, so the construction of outdoor table tennis tables has gradually decreased. From the perspective of site satisfaction, township schools are

generally more satisfied with sports venues than urban schools. Middle schools are more satisfied with sports venues than primary schools, and urban and townships have similar satisfaction with sports equipment. However, the level of satisfaction is higher than that of sports venues, and the reason why urban areas' satisfaction with sports venues and equipment is lower than that of townships is the shortage of school sports venues and equipment in urban areas. Another aspect is that physical education teachers in urban schools have higher requirements for sports venues and equipment than teachers in townships. Figure 8 shows the comparison of the satisfaction of sports venues in urban and rural areas. At this time, according to the real-time evaluation characteristics of genetic algorithm, we need to timely find out the deficiencies of physical education teaching in urban and township primary and secondary schools, optimize physical education teaching resources, and promote the balanced development of public sports.

Figure 9 shows the survey results of physical education teaching in urban and township primary and secondary schools. From this survey, it can be seen that the class-teacher ratio in urban and township schools has not exceeded 60%. The lack of teachers in school physical education reflects that Chinese school physical education is in a low position in school teaching, and the reason why the class-teacher ratio in urban schools is not much different from that in township schools is similar to that of sports venues and facilities. The compliance rate of physical education courses in urban schools is significantly higher than that in township schools. The proportion of urban schools offering morning exercises is smaller than that of township schools, which is mainly due to the relatively few residential schools with convenient transportation in urban areas. The urban schools are better than the township schools in terms of large-scale inter-class activities, and the forms are also more diverse than the township schools. There is little difference between schools in urban and rural in terms of one-hour exercise time per day and the development of extracurricular activities in schools, but the overall compliance rate of one-hour exercise time per day is less than 60%. Judging from the physical fitness test arranged by the school, the implementation rate of physical fitness test in urban schools is better than that in townships, which is also reflected in the recovered data. The myopia rate in urban schools is up to 40%, and the obesity rate is up to 30%. Both urban and township schools are generally able to hold a track and field meeting every school year, and there is not much difference in the development of ball games and other sports competitions.

The main data comparison of sports venues is shown in Figure 10. Although the development achievements of
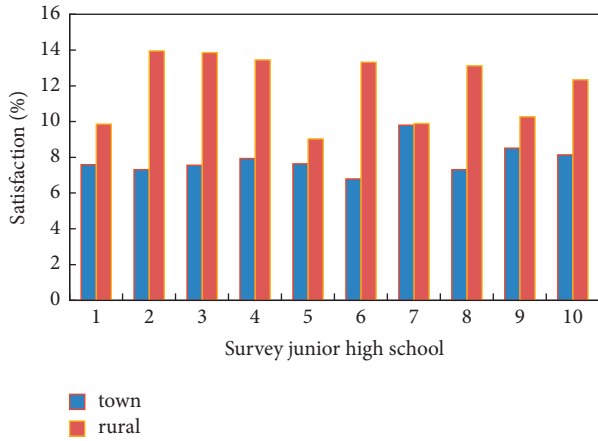
FIGURE 8: Comparison of sports venue equipment satisfaction in urban and rural areas.
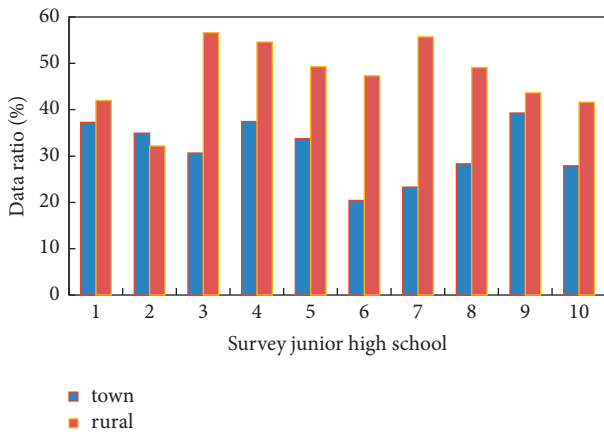
FIGURE 9: Results of a survey on physical education teaching in primary and secondary schools in urban and townships.
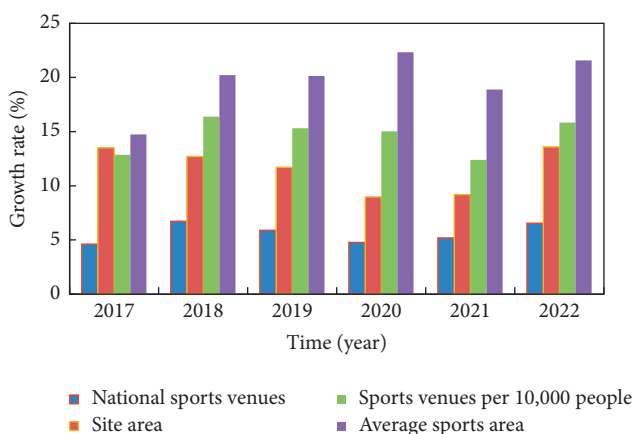
FIGURE 10: Comparison of main data of sports venues.

China's sports public service are huge, so far, the current situation of China's sports public service is not very satisfactory. There are still two flaws: first, the total amount is insufficient. At present, the utilization structure of China's sports public facilities is extremely unreasonable. On the one

hand, the number of stadiums serving competitive sports is huge, and there are tens of thousands of large stadiums for international and domestic competitions, which is second to none in the world. On the other hand, although the abovementioned data illustrating the achievements show that by the end of 2020, there will be 3.7134 million sports venues in the country, but the absolute number of sports venues managed by the education system, government agencies and institutions, and the military system, and the opening rate of these sports venues to the outside world is low.

5. Conclusion

In order to improve the utilization rate and service quality of public sports and balance the contradiction between supply and demand of public sports services, an optimal individual model of the genetic algorithm is established. Excellent individuals can spread rapidly in the area and converge to an optimal solution for the balance between supply and demand of public sports services in a relatively short period of time. At the same time, the algorithm proposed in this paper can effectively achieve the balance of public sports among multiple regions and can focus on optimizing service quality or operating cost according to actual needs. A major problem in the supply of China's sports public services is that the current China's sports public services are seriously insufficient in terms of resource supply. In addition, there is another problem that has attracted much attention in the development of China's sports public services, that is, the "relatively insufficient" supply still exists in regional imbalances and serious imbalances between urban and rural areas. Only by continuously expanding the scale and optimizing the structure in the process of supply, improving the availability and convenience of public sports services for the masses, and achieving the goal of full coupling between supply and demand can public sports serve the whole people in a true sense. The genetic algorithm proposed in this study helps to grasp the whole and each element of public sports service performance by deeply understanding the performance of each structural element of public sports service performance and to clarify the current problems and advantages of public sports service performance. This study makes a statistical analysis of the evaluation results of various structural elements of public sports services and compares the differences in the sub-performance of public sports services between regions through genetic algorithms, in order to promote the improvement and optimization of public sports services. Aiming at many problems existing in the sports public service and some action plans proposed by reference to national policies, this study analyzes the multiple dilemmas in optimizing the sports public service and tries to propose a feasible way to break through the dilemma. The collaborative innovation model of public sports service supply may be more inclined to economically developed regions. For remote and backward areas in China, due to the imperfect market and social organizations, the conditions for coordinated supply are not yet available, so the current development path should still be based on government

support and supply. However, the coordinated supply of public sports services will be the overall trend in the future, and exploring the networked governance model will surely become an important way to carry it. Guided by the needs of residents' sports activities, aiming at the contradiction between supply and demand, a targeted strategy is proposed, and circular monitoring is achieved, so that the intelligent supply of public sports facilities can be realized, the sports interests of residents can be maximized and the informatization of government management can be achieved.

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] G. Xiangwei, "Study on the cooperative innovation model and optimization path analysis of the multiple supply of public sports services," *International Journal for Engineering Modelling*, vol. 31, no. 1, pp. 77–82, 2018.

[2] J. Sun and K. Hu, "The construction of sports public service system for the elderly from the perspective of healthy aging," *Revista Brasileira de Medicina do Esporte*, vol. 27, pp. 66–68, 2021.

[3] D.-H. Chung, "The effect of main performance on the performance of public sports practitioners," *Journal of the Korean society for Wellness*, vol. 12, no. 1, pp. 285–298, 2017.

[4] F. He, "Agricultural climate change based on GIS and remote sensing image and the spatial distribution of sports public services," *Arabian Journal of Geosciences*, vol. 14, no. 11, pp. 1–20, 2021.

[5] Y. Li, "Research on the socialization supply of sports public services under the background in of information age," *Revista de la Facultad de Ingenieria*, vol. 32, no. 12, pp. 16–22, 2017.

[6] R. Tavakkoli-Moghaddam, J. Safari, and F. Sassani, "Reliability optimization of series-parallel systems with a choice of redundancy strategies using a genetic algorithm," *Reliability Engineering & System Safety*, vol. 93, no. 4, pp. 550–556, 2008.

[7] A. Volkanovski, B. Mavko, T. Boševski, and M. ČauševskiČepin, "Genetic algorithm optimisation of the maintenance scheduling of generating units in a power system," *Reliability Engineering & System Safety*, vol. 93, no. 6, pp. 779–789, 2008.

[8] Q. Long and C. Wu, "A hybrid method combining genetic algorithm and Hooke-Jeeves method for constrained global optimization," *Journal of Industrial and Management Optimization*, vol. 10, no. 4, pp. 1279–1296, 2014.

[9] D. W. Gong, J. Sun, and Z. Miao, "A set-based genetic algorithm for interval many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 47–60, 2018.

[10] M. Nemati, M. Braun, and S. Tenbohlen, "Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming," *Applied Energy*, vol. 210, pp. 944–963, 2017.

[11] Y. Yoshitomi, H. Ikenoue, T. Takeba, and S. Tomita, "Genetic algorithm in uncertain environments for solving stochastic programming problem," *Journal of the Operations Research Society of Japan*, vol. 43, no. 2, pp. 266–290, 2000.

[12] C. K. Koffi, H. Djoudi, and D. Gautier, "Landscape diversity and associated coping strategies during food shortage periods: evidence from the Sudano-Sahelian region of Burkina Faso," *Regional Environmental Change*, vol. 17, no. 5, pp. 1369–1380, 2017.

[13] H. S. Schroder, M. M. Yalch, S. Dawood, C. P. Callahan, M. B. Donnellan, and J. S. Moser, "Growth mindset of anxiety buffers the link between stressful life events and psychological distress and coping strategies - ScienceDirect," *Personality and Individual Differences*, vol. 110, no. 1, pp. 23–26, 2017.

[14] F. G. Paes, A. A. Pessoa, and T. Vidal, "A hybrid genetic algorithm with decomposition phases for the Unequal Area Facility Layout Problem," *European Journal of Operational Research*, vol. 256, no. 3, pp. 742–756, 2017.

[15] S. E. Pinar, O. D. Aksoy, and G. Daglar, "Effect of stress management training on depression, stress and coping strategies in pregnant women: a randomised controlled trial," *Journal of Psychosomatic Obstetrics and Gynecology*, vol. 39, no. 3, pp. 1–8, 2017.

Retraction

Retracted: Behavior Data Analysis of English Learners Based on Discrete Dynamic System Modeling 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] H. Chen and Y. Gao, "Behavior Data Analysis of English Learners Based on Discrete Dynamic System Modeling Method," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5409571, 8 pages, 2022.

Research Article

Behavior Data Analysis of English Learners Based on Discrete Dynamic System Modeling Method

Hongli Chen ⓘ and Yuzheng Gao ⓘ

School of Foreign Languages, Handan University, Handan, Hebei 056006, China

Correspondence should be addressed to Yuzheng Gao; yuzhenggao@126.com

With the rapid development of computer science, there are more and more kinds of discrete dynamic systems. Computer integrated system CIMS, network communication database administrator system, and human behavior analysis system are all discrete dynamic systems. At present, many researchers have studied by adding human behavior data to discrete dynamic systems. This paper aims to study the behavior data of English learners by using the discrete dynamic modeling technology of complex systems and the discrete dynamic system modeling method of Petri nets. By adding the behavior data of learners to the discrete dynamic system of fuzzy Petri nets, the system is diagnosed and optimized. The experimental results show that the complex discrete dynamic system in this paper has achieved good experimental results according to the performance indicators selected in theory. Based on the combination of the above technologies and systems, the fuzzy Petri net discrete dynamic system studied in this paper improves the processing speed of English learners' behavior data.

1. Introduction

Discrete dynamic system modeling and data inductive analysis technology is an important direction of data research in recent years [1]. In discrete dynamic system, with the overall development of society, discrete dynamic system and data inductive analysis technology are widely used in medical, financial, heavy industry, education, and other industries [2]. Education industry is the key research direction of many researchers in recent years, mainly by adding the behavior data information generated by the observed learners to the discrete dynamic model for data processing and analysis [3]. The data can be controlled by using discrete dynamic system, and then the data of single type continuous variable can be studied. Through the use of mathematical form and mathematical theory, the modeling, analysis, optimization, and other technologies are systematically improved [4]. The modeling technology of discrete dynamic system under linear constant system has been mature. However, compared with the behavior data processing of learners in the education industry, the ordinary

discrete dynamic modeling technology cannot meet the research needs [5]. The main reason is that the behavior data generated by learners are multifaceted and multiple types of composite data, which is fundamentally different from a single type of variable data. In order to solve the above problems, researchers used Petri net discrete dynamic system, which is mainly used for synchronous, asynchronous, and distributed random data [6]. Petri nets originated from C. A. Petri's doctoral thesis "communication with auto Mata" in 1962. Petri nets are an abstract and formal tool, especially suitable for describing the control flow, competition, concurrency, and asynchronous behavior characteristics of discrete event systems. In recent years, the theory and application of Petri nets have developed rapidly and have been widely used in many fields. The main characteristics of Petri net discrete dynamic system are reachability, flexibility, synchronization, fairness, and so on. After analyzing the behavior data, we describe the data and finally establish the associated state relationship [7]. However, the system also has defects. Under unbounded conditions, there will be state combination collapse and failure to depict

dynamic characteristics. Finally, many researchers solved the above problems by adding fuzzy colored Petri net system. Fuzzy colored Petri net discrete dynamic system can avoid data conflict and combination collapse for behavior information data [8, 9]. Through the data parameters between the data and the model, an automatic fault detection can be carried out, which not only improves the data processing capacity and efficiency but also reduces the system fault rate.

With the rapid development of all walks of life, researchers apply fuzzy Petri net discrete dynamic system to observe human behavior [10]. Fuzzy Petri net discrete dynamic system establishes a fuzzy neural network to express knowledge and self-learning training and finds the internal attributes of the system through fuzzy neural network. After the above "training," the data with adjacent meaning can be neutralized and then changed to the same meaning [11]. The transition conditions of the next data state can be deduced from the existing data state or the State Library set of the next data can be obtained from the known data state [12]. Based on the above content, this paper combines English learners' behavior data with fuzzy Petri net discrete dynamic system and predicts another behavior data state according to one behavior data state. After entering the behavior data state cycle prediction, it is the dynamic process of the whole fuzzy Petri net discrete dynamic system.

This paper is mainly divided into three parts. Section 1 mainly explains the development of Petri net discrete dynamic system. In Section 2, the discrete dynamic modeling technology of fuzzy Petri net is used to combine learners' behavior data. First, fuzzy logic Petri net and neuro fuzzy system are used to model the system, and the behavior data samples of English learners input into the established model are systematically tested. Then, the established Petri net discrete dynamic system is used to detect system faults and further system optimization. Section 3 analyzes the research results of the input English learner behavior data in the Petri net discrete dynamic system and analyzes the model data after fault detection and system optimization of the Petri net discrete dynamic system.

2. The Related Works

Learner behavior data are important data base in Petri net discrete dynamic system [13]. First, the basic structure of fuzzy logic control is constructed by using the fuzzy reasoning algorithm of Petri net to fuzzify the data attributes of learners' various types of behavior data [14]. Dynamic fuzzy Petri net is a good modeling tool for knowledge base system based on fuzzy production rules. It not only combines the graphic description ability of Petri net, makes the representation of knowledge simple and clear, but also reflects the structural characteristics of rules in knowledge base system. It also has the fuzzy reasoning ability of the fuzzy system, which is convenient for the analysis and reasoning of knowledge. It is consistent with the human thinking and cognitive mode and has important value in describing and analyzing the parallel and concurrent behavior of many physical systems and even social systems. The operation of

fuzzy reasoning and data clarity is carried out in the original database, in which the database contains all knowledge in the field of education and the core objectives of control. The database is mainly composed of normal database and fuzzy control rule base [15]. The fuzzy reasoning part is the core part of the data processing of the modeling system. The role of data clarity is to convert the fuzzy processed learner behavior integration data into the final practical application clarity [16, 17]. Many researchers have shown that fuzzy control is an important and effective means to realize English learners' behavior analysis. This technology can fuzzily assimilate many kinds of behavior data. Compared with the traditional discrete dynamic modeling technology, it can more accurately infer and transform different data [18]. The emergence of this technology also greatly improves the accuracy of data processing.

In Germany, the complex discrete dynamic system is applied to detect the types of TV viewers like to watch [19]. Through this technology, the viewing type, time, viewing times, and other data of each user when watching TV can be systematically integrated. The use of this technology greatly reduces the labor cost of human door-to-door visits. Through a period of observation, the test results can be easily obtained.

In the United States, the complex discrete dynamic model is applied to the logistics industry [20]. As we all know, the United States is a major importer and exporter, and its economic industry has always ranked first in the world. The demand for the logistics industry in the United States is large. At the same time, it is necessary to meet the requirements of efficient and reliable reduction of goods flow and goods reserve [21]. Based on the above two points, by adding the model, the simplification can be reflected in the process of modeling the whole logistics. Finally, the rapid and stable pipeline application of logistics is realized.

China's complex discrete dynamic systems are mainly used in machinery manufacturing industry. China's development from the electric steam era to today's automatic machinery production is absolutely inseparable from discrete modeling technology [22]. Through the reference of this technology, the data generated in the manufacturing process of mechanical products can be clearly seen. Technicians can observe the data and finally confirm the performance of mechanical products.

The UK traffic management industry uses complex discrete dynamic systems. The time control of traffic lights is mainly applied by observing the road congestion on each section of the road [10]. The overall analysis of the data can be generated by converting the monitoring video data to the discrete system. Finally, the monitoring administrator regulates the traffic light time of each section through the data trend, which greatly avoids traffic congestion and accidents. According to the above application of complex network dynamic modeling technology in various countries, this paper systematically studies the behavior data samples generated by English learners based on Petri discrete dynamic system [23]. In the discrete dynamic system, the main part is to input the behavior data uniformly and judge the overall performance of the system through the output

results. Finally, the whole Petri discrete dynamic system is further improved by adding optimization and fault detection technology [24].

3. Research on English Learners' Behavior Data Based on Petri Discrete Dynamic Modeling Technology

3.1. Modeling of English Learners' Behavior Data Based on Petri Discrete Dynamic System. With the development of science and technology, artificial neural network has developed more and more rapidly in recent years. Neural network is mainly similar to the ability of learning and organization in human brain. It also has a neural layer that transmits different levels of information. It can also feedback system information and process information in real time. The Petri net structure of discrete dynamic system is complex, but as long as we make good use of the essence of the system, we can flexibly use the system to mine the behavior data of English learners. First, it has a precise definition and a solid mathematical foundation. Compared with many other nonformal block diagram technologies, it avoids ambiguity, uncertainty, and contradiction. Second, this formal system can be used to reflect on the process, such as establishing a specific pattern, and also promote the use of many analysis technologies (such as the technology of analyzing performance and the technology of verifying logical properties). No matter how complicated the data are, it can be processed dynamically under Petri net. This paper establishes a fuzzy neural network, which can learn and transform data by itself through two key points: database and data change [25]. Through fuzzy neural network, we can explore the internal attributes of data and then add the means of learning and training, so that we can homogenize the learning behavior data. This cycle constitutes the whole automatic input data processing process as shown in Figure 1.

It can be seen from Figure 1 that fuzzy Petri net can automatically fuzzify the data when processing learning behavior data. By putting the data into the database and further corresponding the fuzzy operation parameters with the total data of the database, the purpose of fuzzy data is achieved. The design of fuzzy Petri net discrete dynamic system is based on the fuzzy set theory proposed by Zadeh. After adding two members, the overall mathematical formula is as follows:

$$\mu_A(x) = \begin{cases} 1 \ if \ x \in X \\ 0 \ if \ X \notin x \end{cases}, \tag{1}$$

X mainly represents fuzzy data set. Fuzzy data processing system and neural network are added to Petri net, and finally, fuzzy Petri net discrete dynamic system is generated. The seven-element set in the system is as follows:

$$FENP = \{P, T, F, C, K, \gamma, \delta\}, \tag{2}$$

where P is the data set of the database and T is the data set of the fuzzy transition of the whole data. F is the set of data reflection arcs in the system, and C is the dynamic state of
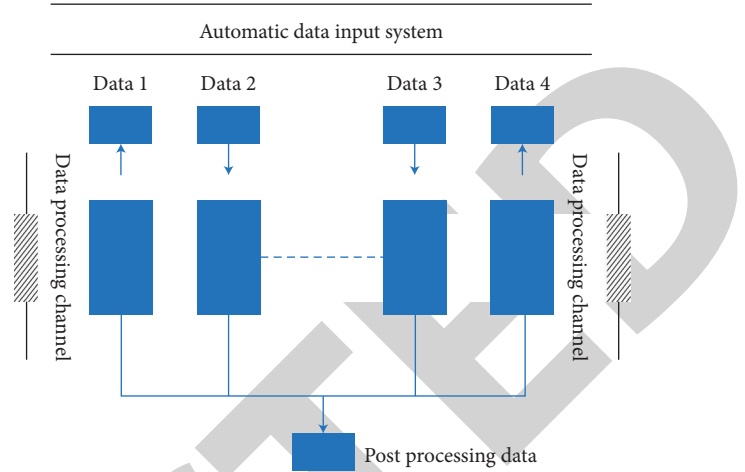
Figure 1: Automatic input data processing flowchart.

data hiding and output. K is a finite set of fuzziness. The model of fuzzy Petri net can be obtained only after modeling the discrete dynamic system of fuzzy Petri net. The specific modeling flowchart is shown in Figure 2.

It can be seen from Figure 2 that the whole process of modeling the discrete dynamic system based on Fuzzy Petri net, after fuzzifying, neural networking, and coloring the added learner behavior data, carry out an overall circular operation, and finally carry out the state prediction and model establishment of a system. The total number of network data nodes and the applicability of fuzzy rules are the following equation:

$$N_2 = \sum_{i=1}^{n} m_i, \tag{3}$$
$$\alpha_j = \min\left\{\mu_1^{i_1}, \mu_2^{i_2}, ..., \mu_n^{i_n}\right\} \text{或} \alpha_j = \mu_1^{i_1}, \mu_2^{i_2}, ..., \mu_n^{i_n}.$$

It can be seen from the above equation that the function of fuzzy rules is to match the hard conditions of data. After calculating the applicability of the rule, the applicability is approximated by a local network. Finally, through the normalization algorithm, the algorithm formula is as follows:

$$\overline{\alpha_j} = \frac{\alpha_j}{\sum_{i=1}^{m} \alpha_i}, i = 1, 2, ..., m, \tag{4}$$

Through the output data generated by the network data, the subsequent data cycle rules of the fuzzy network are calculated. Based on the above formula, the fuzzy Petri net discrete dynamic system is obtained, and the network diagram of the system is shown in Figure 3.

As can be seen from Figure 3, the input learner behavior data is a set of databases in the fuzzy Petri net discrete dynamic system. Through the input of data, the behavior data are fuzzified at a single point, and finally the fuzzified data are collected. The connection weight of data connection and the center value and width of data shall be calculated. The connection weight algorithm formula is as follows:

$$y = \frac{\sum_{j=1}^{m} \alpha_j y_j}{\sum_{j=1}^{m} \alpha_j} = \sum_{j=1}^{m} \overline{\alpha_j} y_j. \tag{5}$$
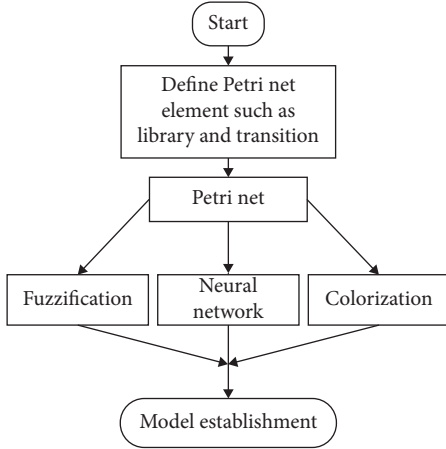
FIGURE 2: Modeling flowchart.

According to the above connection weight algorithm formula, if the fuzzy division number of the input data is limited, the final connection value can be obtained through the error function combined with the function of the data node. The formula is as follows:

$$E = \frac{1}{2}\sum_{i=1}^{r}(t_i - y_i)^2,$$

$$p_{ji}^{k}(l+1) = p_{ji}^{k}(l) + \beta(t_k - y_k)\overline{\alpha_j}x_i,$$

(6)

After obtaining the center value and width value, the modeling part of the whole fuzzy Petri net discrete dynamic system is completed. The calculation formula of center value and width is as follows:

$$\frac{\partial E}{\partial c_{ij}} = -\delta_{ij}^{(2)}\frac{2(x_i - c_{ij})}{\sigma_{ij}^2},$$

$$\frac{\partial E}{\partial c_{ij}} = -\delta_{ij}^{(2)}\frac{2(x_i - c_{ij})^2}{\sigma_{ij}^3},$$

$$c_{ij}(k+1) = c_{ij}(k) - \beta\frac{\partial E}{\partial c_{ij}},$$

$$\sigma_{ij}(k+1) = \sigma_{ij}(k) - \beta\frac{\partial E}{\partial c_{ij}}.$$

(7)

Through the above algorithm of data connection weight, center value, and width value, the function of the whole fuzzy Petri net discrete dynamic system is improved. After adding the behavior data of English learners to the system, the network test is carried out, and the prediction database is added for comparison, as shown in Figure 4.

It can be seen from Figure 4 that by comparing the input database and output database of data with the prediction database, the input database and output database correspond to each other in the system test results, and the ultimate goal of establishing the discrete dynamic model of fuzzy Petri net is achieved [26]. It can also quickly find the excitation change conditions that meet the data fuzziness and realize the automatic data processing.
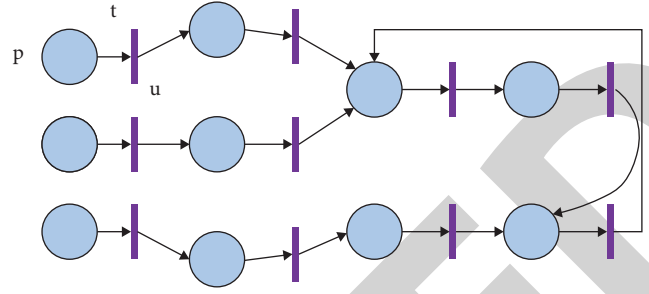
FIGURE 3: Fuzzy Petri net graph.
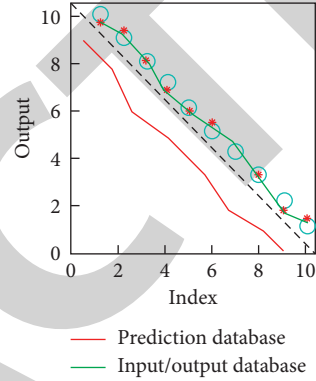
FIGURE 4: Network test node diagram.

3.2. Research on Fault Detection and Optimization of English Learners' Behavior Data Based on Petri Discrete Dynamic System. In the fuzzy Petri net discrete dynamic system, it is important to detect and eliminate the faults in the system in time. For the discrete dynamics of fuzzy Petri nets studied in this paper, the complexity of the system is considered. It mainly adopts the backward tree reasoning method, which can quickly find out the faults in the running system and eliminate them accurately. After adding the backward reasoning method, we first learn the learner behavior data in the fuzzy Petri net discrete dynamic system and then make all changes of all relevant data in the database. Finally, repeat the cycle operation. In the establishment of backward tree, the main condition is that its database of all input data, China's membership function cannot be an empty set, and must be greater than the basic set value. In this paper, with the help of computer, the fuzzy Petri net is transformed into a backward tree, and then the behavior sample data of English learners are input into the fuzzy Petri net discrete dynamic system which can automatically diagnose faults. The fault rate diagnosis trend chart is obtained, as shown in Figure 5.

It can be seen from Figure 5 that the fuzzy Petri net discrete dynamic system with backward tree reasoning method greatly improves the efficiency of fault diagnosis compared with the ordinary discrete dynamic system without adding. Through the addition of backward tree reasoning, the sample data processing of learners' behavior can also deal with the faults. Then, the error of data processing in the system is greatly reduced.
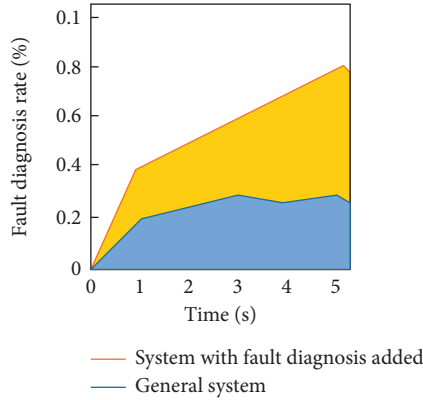
Figure 5: Trend chart of fault diagnosis rate.

The above content is to add the ability of fault detection and fault handling to the system. In addition to the fault handling ability, it is necessary to optimize the performance of the whole fuzzy Petri net discrete dynamic system. In this paper, the dynamic programming method is used to improve the performance optimization of the whole system, and the maximum and minimum algebra methods are used to stimulate the data in the database and construct the functional relationship [27]. Genetic programming (GP) is developed on the basis of genetic algorithm. It can combine structure estimation with parameter estimation and can establish a system dynamic model by using observation data without complete model structure information. It solves the shortcomings of genetic algorithm that cannot describe hierarchical problems and lacks dynamic variability, and provides a new idea and method for complex system modeling. It is more and more widely used in dynamic system modeling. The application of genetic programming in discrete dynamic system modeling is to solve the modeling of the structure and correlation function of discrete dynamic system. With the use of the two methods, a good system control theory is added to the study of learners' writing analysis by Petri net discrete dynamic system, and the research idea is further expanded. The constructor formula of max and min algebraic method is as follows:

$$t_i(n) = [00d_i000][t_i(n-1)] \oplus [0000d_n''0],$$
$$\begin{cases} T(n) = \varphi(T_i(n-1) + c(n)) \\ c(0) = c_0 \end{cases}. \qquad (8)$$

In the constructor of Max and min algebra, the purpose is to make the real-time data change by delaying all databases. Once the data change, the system can be quickly activated, and then the data will be transferred to the output database, which greatly enhances the rate of data transfer. Through the above maximum and minimum algebraic methods, the constructor formula can only optimize a small part of the system, and the dynamic programming method is also needed to complete the final optimization. The comparison chart of cyclic data processing rate under the two algorithms is shown in Figure 6.

As can be seen from Figure 7, by adding the above two algorithms, the delay to the database in the fuzzy Petri net discrete dynamic system obtains the best effective

performance. In the dynamic programming method, in order to obtain the cycle optimal solution of the system studied in this paper, we must first find the performance index. The formula of performance index is as follows:

$$J_N = \sum_{i=1}^{N-1} \varphi_0(t(n), c(n)). \qquad (9)$$

By introducing the formula into the maximum and minimum algebraic methods, the following formula is obtained:

$$\begin{aligned} J_N &= \varphi_0(t(0), c(0)) + \varphi_0(t(1), c(1)) + \dots \dots \\ &= \varphi_0(t(0), c(0)) + \varphi_0(t(0), C(0), c(1)) + \dots \dots \end{aligned}. \qquad (10)$$

From the obtained formula, we can see that J_N only depends on the data behind the data node. Once the node data of optimal control is obtained, the minimum value of J_N depends on the initial data conditions. The expression of the optimal control quantity is as follows:

$$J_N = J_N(t(0), C(0), c(1), ..., c(N-1))$$
$$c_N^*(t(0)) = \min_{c(0),...,c(N-1)} \{(J_N(t(0)t, nCq(0), c(1), ..., cN-1)\}. \qquad (11)$$

Through the expressions of the above two algorithms, the performance of fuzzy Petri net discrete dynamic modeling is systematically optimized. Based on the optimized system, the performance of the optimized fuzzy Petri net discrete dynamic system and the ordinary Petri net discrete dynamic system is compared by adding English behavior sample data, as shown in Figure 7.

As can be seen from Figure 7, the data processing performance and efficiency of the fuzzy Petri net discrete dynamic system with the addition of the maximum and minimum algebra method and the dynamic programming method are high. Although the traditional ordinary Petri net discrete dynamic system has good processing efficiency, the system combining the two algorithms can better process the behavior data of English learners. For the simplification of complex formulas and customization of balance standards, better optimization methods may appear, which need further research in the future.

4. Analysis of Research Results of English Learners' Behavior Data Based on Petri Discrete Dynamic Modeling Technology

4.1. Analysis of Research Results of System Modeling Based on Petri Discrete Dynamic System and English Learners' Behavior Data. After modeling the fuzzy Petri net discrete dynamic system, this paper puts the behavior data of English teaching learners into the fuzzy Petri net discrete dynamic system. In this experiment, more than 3000 groups of experimental sample data with different attributes were prepared. After adding the learner behavior data, the sample data fuzzizes the input sample data with different attributes in the system and then integrates and outputs the fuzzy data. In order to more clearly see the data processing rate of learner behavior
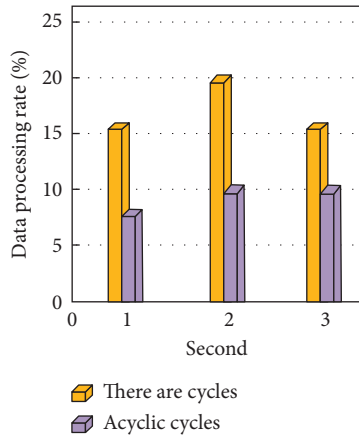
Figure 6: Comparison chart of data processing rate under cycle.
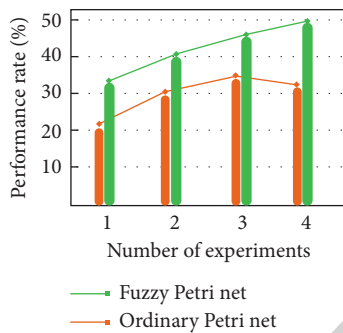
Figure 7: Comparison of performance rate between fuzzy Petri net and ordinary Petri net.

Figure 8: Waveform generated by discrete dynamic system with sample data.

Figure 9: Comparison of data processing rate of three kinds of discrete dynamic systems.

sample data in complex discrete dynamic system, the sample data are also put into three different types of discrete dynamic systems, namely ordinary discrete dynamic system, Petri net discrete dynamic system, and fuzzy Petri net discrete dynamic system. Through the result data diagrams generated by different kinds of complex discrete dynamic systems, we can better judge whether the fuzzy Petri net discrete dynamic modeling technology studied in this paper is qualified. It can also briefly reflect the ability gap of different types of complex discrete dynamic systems in dealing with English learners' behavior data. The waveform change generated by the sample data discrete dynamic system is shown in Figure 8. The comparison of data processing capabilities of three different types of complex discrete dynamic systems is shown in Figure 9.

It can be seen from Figure 8 that the waveform change curve generated by the discrete dynamic system of sample data is similar to the normalized number curve. Because the fuzzy Petri net discrete dynamic model studied in this paper is a stable structure, it will not change greatly without modifying the system. The internal data of the system studied in this paper can be "self-learning" and data fuzzification, so the data waveform trend is similar to the standard. It can be seen from Figure 9 that among the three different types of complex discrete dynamic models, the data processing capacity of the fuzzy Petri net discrete dynamic system studied in this paper is high speed and stable. It will

not cause the problems of low data processing efficiency and unstable processing rate compared with the other two systems. According to the above conclusions, the fuzzy Petri net discrete dynamic model has good performance in the processing of English learners' sample data.

4.2. Analysis of Fault Detection and Optimization Based on Petri Discrete Dynamic System.

Discrete event dynamic system is different from continuous variable dynamic system, and Petri net is one of the effective means to model and analyze discrete event dynamic system. Its main purpose is to analyze the working performance of the system, that is, to optimize the system under certain conditions, so that the system can provide services more effectively and provide theoretical basis for designing new systems or improving existing systems. Then, the system is driven to work according to people's expectations through external input. Petri net is bounded, so the reachable data identification is also bounded, which also means that the traversal of the system exists. In order to avoid the deadlock problem caused by the detection system in processing sample data, the deadlock problem is ignored in this experiment. In order to further verify, two technical performances of fault processing and data optimization processing are added to the fuzzy Petri net discrete dynamic system. In this experiment, a large number of learner behavior data with different attributes are processed by the optimized fuzzy Petri net discrete dynamic system. In order to make the experimental
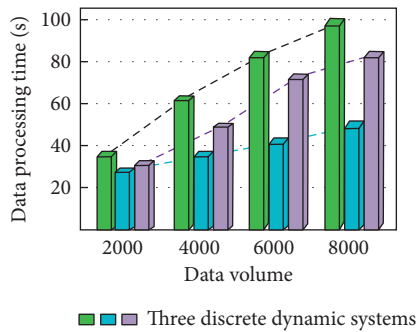
Figure 10: Processing time distribution of three discrete dynamic systems under different data quantities.

results more comparable, ordinary discrete dynamic system and Petri net discrete dynamic system are added to compare the experimental results. This paper makes a comparative study on the processing time of learner behavior data by three systems through four groups of different data. The final experimental results are shown in Figure 10.

It can be seen from Figure 10 that among the three complex discrete dynamic systems, the optimized fuzzy Petri net discrete dynamic system has the least time to process sample data. The second is the nonoptimized fuzzy Petri net discrete dynamic system, and the ordinary discrete dynamic system has the longest processing time. And in the case of the rapid growth of data, the optimized fuzzy Petri net discrete dynamic system still maintains high efficiency and high quality. Therefore, through the experimental results of this paper, the optimized fuzzy Petri net discrete dynamic system is suitable for the analysis of English learners' behavior data in the field of education.

5. Conclusion

The modeling technology of fuzzy Petri net discrete dynamic system is based on the advantages of concurrency time and graphical representation in its essence. Based on the research status of this type at home and abroad, this paper further studies Petri net discrete dynamic systems. Fuzzy technology is mainly added to Petri net to further establish the discrete dynamic model of fuzzy Petri net. The whole experimental research is based on fuzzy control technology and then supported by the theoretical technology in modeling. The traditional discrete modeling technology cannot meet the data synchronous processing under massive data. In order to achieve the experimental purpose, the minimum and maximum algebra method and dynamic programming method are proposed to optimize the data cycle of the system. It also adds fault diagnosis and processing functions to the original system by combining the principle of "backward tree." The results show that the fuzzy Petri net discrete dynamic system with the above functions is easier to understand and also provides a new research idea for the optimization of system performance. The main reason for the success of this experiment is that it simplifies the complexity of the system and is easier to be applied in practice. The selection of optimization criteria is of great significance in the optimization process of Petri nets and

complex discrete dynamic systems. The experimental results show that the complex discrete dynamic system in this paper has achieved good experimental results based on the selection of performance indexes under theory. Based on the combination of the above technologies and systems, the fuzzy Petri net discrete dynamic system studied in this paper improves the processing speed of English learners' behavior data.

Data Availability

The figures used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] L. Shen, L. G. Di, De Zhu et al., "Mechanical responses of Steel fiber–reinforced concrete after Exposure to high Temperature: Experiments and Mesoscale discrete modeling," *Journal of Engineering Mechanics*, vol. 147, no. 11, Article ID 04021084, 2021.

[2] Z. Chen and X. Li, "Designing corridor systems with modular autonomous vehicles enabling station-wise docking: discrete modeling method," *Transportation Research Part E*, vol. 152, Article ID 102388, 2021.

[3] T. Loossens, F. Tuerlinckx, and S. Verdonck, "A comparison of continuous and discrete time modeling of affective processes in terms of predictive accuracy," *Scientific Reports*, vol. 11, no. 1, p. 6218, 2021.

[4] Y. Feng, "A generic energy conserving discrete element modeling strategy for concave particles represented by surface triangular meshes," *International Journal for Numerical Methods in Engineering*, vol. 122, no. 10, pp. 2581–2597, 2021.

[5] G. Binda, A. Pozzi, D. Spanu, F. Livio, S. Trotta, and R. Bitonte, "Integration of photogrammetry from unmanned aerial vehicles, field measurements and discrete fracture network modeling to understand groundwater flow in remote settings: test and comparison with geochemical markers in an Alpine catchment," *Hydrogeology Journal*, vol. 29, no. 3, pp. 1203–1218, 2021.

[6] J. C. Quezada and G. E. Villavicencio, "Discrete modeling of waste rock dumps stability under seismic loading," *EPJ Web of Conferences*, vol. 249, Article ID 11013, 2021.

[7] M. Fukushi, C. A. Guevara, and S. Maldonado, "A discrete choice modeling approach to measure susceptibility and subjective valuation of the decoy effect, with an application to route choice," *Journal of Choice Modelling*, vol. 38, Article ID 100256, 2021.

[8] G. Min, D. Fukuda, S. Oh et al., "Three-Dimensional combined finite-discrete element modeling of shear fracture process in Direct Shearing of rough concrete–rock Joints," *Applied Sciences*, vol. 10, no. 22, p. 8033, 2020.

[9] B. Khajji, A. Labzai, A. Kouidere, O. Balatif, M. Rachik, and F. Simões, "A discrete mathematical modeling of the influence of Alcohol Treatment centers on the Drinking dynamics using optimal control," *Journal of Applied Mathematics*, vol. 2020, pp. 1–13, 2020.

[10] J. Zhang, J. Eisenträger, S. Duczek, and C. Song, "Discrete modeling of fiber reinforced composites using the scaled boundary finite element method," *Composite Structures*, vol. 235, Article ID 111744, 2020.

Retraction

Retracted: Assessment of Ecological Environment Quality for Urban Sustainable Development Based on AHP

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. Zhou, "Assessment of Ecological Environment Quality for Urban Sustainable Development Based on AHP," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4056713, 9 pages, 2022.

Research Article

Assessment of Ecological Environment Quality for Urban Sustainable Development Based on AHP

Liang Zhou [iD]

School of Design and Art, Hunan Institute of Technology, Hengyang 421002, Hunan, China

Correspondence should be addressed to Liang Zhou; 2015001977@hnit.edu.cn

Cities are gradually developed on the basis of adapting and transforming the natural environment. In a certain urban area, human activities, natural environment, and other factors and their mutual influence constitute the urban ecological environment. Therefore, the evaluation of urban ecological environment quality is of great significance to the analysis of urban development. This paper takes a city in Western China as the evaluation object, uses AHP to determine the index weight, reasonably analyzes the current situation of the urban ecological environment, and further comprehensively evaluates the quality of the urban ecological environment. The study shows that from 2013 to 2018, the comprehensive capacity of the city's ecological environment quality showed a steady upward trend, except that the natural disasters of floods and mudslides in 2014 had a certain degree of fluctuation. The comprehensive index of ecological environment quality has increased from 0.337 in 2013 to 0.412 in 2018. The overall level is still low, but the development speed is relatively stable. The urban ecological environment has been gradually improved, and society, economy, and nature have maintained a certain degree of sustainable development.

1. Introduction

The product of the development of human society to a certain stage has produced cities, so the emergence of cities is a symbol of human progress. In 1971, UNESCO put forward the concept of "eco city" in the man and biosphere plan, emphasizing the realization of a complex natural symbiosis system and sustainable development of social economy, and the creation of harmony between man and nature. However, the harmonious coexistence between man and nature has not been realized. On the contrary, the greenhouse effect, global warming, and other phenomena indicate that the Earth on which civilization depends is suffering damage. The current large-scale urbanization and urban modernization in China are facing severe ecological problems: first, the natural resources and the environment are severely damaged [1]. The city is an artificial system based on the natural environment. If this artificial system is in the process of establishment and development, inadequate treatment of nature and protection of nature can easily cause the original ecological environment in the urban area to gradually shrink

and trigger a vicious ecological environment effect [2, 3]. The second is that the living resources are seriously polluted. The water pollution caused the shortage of water resources and air pollution has become a bottleneck in the development of many cities. The third is soil heavy metal pollution, which occupies more and more solid wastes [4–6]. With the continuous improvement of the output requirements of agricultural products per unit area, there are more and more chemical fertilizers and pesticides per unit of land, and more and more undecomposable substances remain in the soil. Once the soil is polluted, it is difficult to control because there are many elements of the urban ecological environment. An urban ecosystem is an open ecosystem with large capacity, multiflow, high density, and fast operation. The imperfect relationship network of the system and the intensity of various flows in the natural ecosystem make its decomposition function inadequate and its self-regulation and self-sustaining ability weak. Besides environmental pollution factors, many scholars also use the status of natural resource holdings and biodiversity to reflect the quality of the urban ecological environment [7, 8]. Broadly speaking,

the quality of the urban ecological environment is a combination of the quality of the socio-economic environment and the quality of the city's natural environment. From this perspective, the quality of the urban ecological environment can include several aspects such as air quality, acoustic environmental quality, solid waste, soil environmental quality, and water environmental quality [9]. As the concept of sustainable development gradually gains popularity among the people, many countries regard sustainable development as the most important concept and principle when evaluating regional ecological quality and put forward a number of index systems that are conducive to the sustainable development of cities [10–12]. The so-called sustainable development capacity of the ecological environment is the dynamic identification of the total capacity of the regional environment. The human development of the region, the utilization of resources, and the transformation of nature should be maintained within the allowable capacity of the environment. In other words, only by maintaining the quality of the real environment within the allowable bearing standard can we seek more reasonable development. The sustainable development ability is the necessary guarantee for the smooth implementation of the sustainable development strategy. Specifically, sustainable development capacity involves decision-making, management, economy, resources, science and technology, human resources, and other aspects of a certain region. Local governments are the main force in implementing the sustainable development strategy. For example, in 1990, the United Nations Organization for Economic Cooperation and Development (OECD) launched a project to study ecological environment indicators and adopted the PSR system, which mainly represents the pressure caused by human activities on the ecological environment [13–15]. The United States Environmental Protection Agency, Washington Development and Research Office, and other units jointly conducted a comprehensive evaluation of the ecological health of Atlanta in the United States. They selected 32 indicators as the evaluation index system, including population per unit area, rate of population change, human utilization Index, road density index, annual deposition of nitrogen in a certain period of time, annual deposition of sulfur in a certain period of time, annual rate of change of ozone in a certain period of time, and the ratio of elements such as rivers and forests in urban areas [16–18].

The implementation of urban community planning and design in China inevitably requires saving resources and harmonious coexistence with nature, and the construction of ecological civilization is its due meaning. The basis for strengthening the construction of ecological civilization is to systematically understand the evaluation system and risk warning of ecological civilization construction. Only by establishing the evaluation system of ecological civilization construction as soon as possible can it play a due role in the management department's formulation of corresponding supervision strategies, the establishment of its key monitoring areas, the improvement of its ecological supervision efficiency, and the prevention of ecological risks. This article is based on ecological theory and sustainable development theory. This study takes a city in western China as an example and uses the analytic hierarchy process to establish an urban ecological environment quality evaluation index system to provide guidance for the analysis and evaluation of urban ecological environment quality.

The evaluation of urban ecological environment quality needs to be considered from both qualitative and quantitative aspects, analyze the factors affecting the urban ecological environment, understand the relationship between urban ecological environment quality and urban economic development and social stability, and comprehensively consider various factors in the evaluation of urban ecological environment construction, so as to better complete the evaluation and analysis of urban ecological environment quality.

2. Methods and Materials

2.1. AHP Model. Analytic hierarchy process (AHP) is a systematic method that takes a complex multiobjective decision-making problem as a system, decomposes the objective into multiple objectives or criteria, and then decomposes it into several levels of multiple indicators (or criteria and constraints), and calculates the single ranking (weight) and total ranking of the levels through the qualitative index fuzzy quantification method, so as to be the objective (multiobjective) and multischeme optimization decision-making. The basic principle of the analytic hierarchy process is to divide the basic elements of the scheme into layers and evaluate each layer separately. Each layer needs to make a judgment and comparison on the basis of the upper layer to finally form good calculation elements so that relevant weights and evaluation criteria can form the best evaluation scheme, and the derivation of the best scheme is carried out by means of weighted average.

The target layer is the comprehensive index of the ecological environment quality of the city. The subtarget layer includes social, economic, and natural indicators. The criterion layer includes population factors, social security, resource allocation, pollution control, industrial structure, economic income, air quality, acoustic environment quality, water environment quality, and biological environment quality [19]. Indicators include natural population growth rate, urbanization level, number of hospital beds per 10,000 people, education investment as a percentage of GDP, per capita housing area, urban population registered unemployment rate, industrial wastewater discharge compliance rate, the comprehensive utilization rate of industrial solid waste, harmless treatment rate of domestic waste, the proportion of tertiary industry in GDP, GDP growth rate, annual disposable income per capita, GDP per capita, urban-rural income ratio, smoke and dust control area coverage, urban air pollution index API, the average value of the environmental noise equivalent sound level in the urban area, the compliance rate of urban drinking water source water quality, the urban green coverage rate, and the urban per capita public green area. The analytic hierarchy process model is shown in Table 1.

Table 1: The analytic hierarchy process model of urban ecological environment quality evaluation.

Target layer	Subtarget layer	Criterion layer	Index layer
Comprehensive index of ecological environment quality, A	Natural indicators, B3	Biological environment quality, C10	Per capita public green area, D20
			Urban green coverage rate, D19
		Water environmental quality, C9	Drinking water source water quality compliance rate, D18
		Acoustic environmental quality, C8	Average value of environmental noise equivalent sound level, D17
		Air quality, C7	Air pollution index, D16
			Coverage rate of smoke and dust control area, D15
	Economic indicators, B2	Economic income, C6	Urban-rural income ratio, D14
			GDP per capita, D13
			Annual disposable income per capita, D12
		Industrial structure, C5	GDP growth rate, D11
			The proportion of tertiary industry in GDP, D10
	Social indicators, B1	Pollution control, C4	Harmless treatment rate of domestic garbage, D9
			Comprehensive utilization rate of industrial solid waste, D8
			Industrial wastewater discharge compliance rate, D7
		Resource allocation, C3	Registered unemployment rate of urban population, D6
			Housing area per capita, D5
		Social security, C2	Education investment as a proportion of GDP, D4
			Number of hospital beds per 10,000 people, D3
		Demographic factors, C1	Urbanization level, D2
			Natural population growth rate, D1

2.2. Analytic Hierarchy Process Calculation. The judgment matrix between the target layer and the subtarget layer is shown in Table 2.

The judgment matrix between the subtarget layer and the criterion layer is shown in Table 3–5.

The weight of the index system is shown in Table 6.

2.3. Comprehensive Evaluation Calculation. This article shows the relevant data of a city in western China from 2013 to 2018. The analytic hierarchy process is used to calculate the weight of each indicator, and on this basis, the corresponding indicator values of each level in the city's urban ecological environment quality evaluation indicator system are calculated. Finally, the indicator value of the target layer is calculated, which is the final evaluation of the comprehensive index of urban ecological environment quality in this city.

2.3.1. Calculation of the Evaluation Index at the Index Level.

When the index value is as large as possible, $Di = \dfrac{Zi}{Si}$,

When the index value is as small as possible, $Di = \dfrac{Si}{Zi}$. (1)

Here, Di is the evaluation index value of the i-th index, Zi is the data value of the i-th index, and Si is the standard value of the i-th index.

2.3.2. Calculation of the Criterion-Level Evaluation Index. The criterion-level evaluation index is derived from the arithmetic average value of the evaluation index of the index-level indicators contained in the level, and its calculation formula is as follows:

$$Ci = \sum_{i=1}^{m} \frac{Di}{m}. \tag{2}$$

Here, Ci is the evaluation index value of the i-th index, m is the number of index-level indexes included in the criterion level, and Di is the evaluation index value of the i-th index in the index level.

2.3.3. Calculation of Subtarget Layer Index. The evaluation index of the subobjective level is obtained by multiplying the evaluation index of the criterion level index contained in it by their respective weights and summing them. The calculation formula is as follows:

$$Bi = \sum_{i=1}^{n} Ci \cdot Wi. \tag{3}$$

TABLE 2: Judgment matrix between the target layer and the subtarget layer.

Comprehensive index of ecological environment quality, A	Social indicators, B1	Economic indicators, B2	Natural index, B3
Social indicators, B1	1	1	1/3
Economic indicators, B2	1	1	1/3
Natural index, B3	3	3	1

CI: RI < 0.1, with satisfactory consistency.

TABLE 3: Judgment matrix of three indicators of a social environment.

Social environmental indicators, B1	Demographic factors, C1	Social security, C2	Resource allocation, C3	Pollution control, C4
Demographic factors, C1	1	1/5	1/3	1/5
Social security, C2	5	1	1	1/3
Resource allocation, C3	3	1	1	1/3
Pollution control, C4	5	3	3	1

CI: RI < 0.1, with satisfactory consistency.

TABLE 4: Judgment matrix of two indicators of the economic environment.

Economic and environmental indicators, B2	Industrial structure, C5	Economic income, C6
Industrial structure, C5	1	1
Economic income, C6	1	1

CI: RI < 0.1, with satisfactory consistency.

TABLE 5: Judgment matrix of two indicators of the natural environment.

Natural indicators, B3	Air quality, C7	Acoustic environmental quality, C8	Water environmental quality, C9	Biological environment quality, C10
Air quality, C7	1	3	1/7	1/5
Acoustic environmental quality, C8	1/3	1	1/9	1/7
Water environmental quality, C9	7	9	1	1/3
Biological environment quality, C10	5	7	3	1

CI: RI < 0.1, with satisfactory consistency.

Here, Bi is the evaluation index value of the i-th index, n is the number of criterion level index items included in the subobjective layer, Ci is the evaluation index value of the i-th index of the criterion level, and Wi is the weight of the i-th index of the criterion level.

2.3.4. *Calculation of the Target Layer Index.* The comprehensive index of the target layer is based on the indexes of each layer and is added again according to their respective weights. The calculation formula is as follows:

$$A = \sum_{i=1}^{n} Bi \cdot Wi, \qquad (4)$$

where A is the comprehensive evaluation index value of the target layer, n is the number of subtarget layer index items included in the target layer, Bi is the evaluation index value of the i-th index of the subtarget layer, Wi is weight of the i-th index of the subtarget layer.

2.4. *Evaluation Basis.* The numerical value of the comprehensive index calculated according to the calculation does not actually have any image significance. It is necessary to classify and define the comprehensive index of urban ecological environment quality. In the grade standard, the suitability index is equal to 0.65 as the passing line. The larger the value, the higher the suitability grade. It also shows that the quality of the urban ecological environment is better, and the sustainable development is stronger. Conversely, the smaller the value, the lower the suitability level, and it also indicates that the quality of the urban ecological environment is worse, and the sustainable development is weaker (Table 7).

3. Result

3.1. *Analysis of Comprehensive Index of Ecological Environment Quality.* From 2013 to 2018, the comprehensive capacity of the city's ecological environment quality generally showed a steady increase. Except for the summer floods and

TABLE 6: Index system weight.

Subtarget layer		Criterion layer		Index layer	
Indicator	Weight	Indicator	Weight	Indicator	Weight
Natural indicators, B3	0.6	Biological environment quality, C10	0.312	Per capita public green area, D20	0.156
				Urban green coverage rate, D19	0.156
		Water environmental quality, C9	0.209	Drinking water source water quality compliance rate, D18	0.209
		Acoustic environmental quality, C8	0.026	Average value of environmental noise equivalent sound level, D17	0.026
		Air quality, C7	0.053	Air pollution index, D16	0.027
				Coverage rate of smoke and dust control area, D15	0.027
Economic indicators, B2	0.2	Economic income, C6	0.1	Urban-rural income ratio, D14	0.018
				GDP per capita, D13	0.037
				Annual disposable income per capita, D12	0.045
		Industrial structure, C5	0.1	GDP growth rate, D11	0.025
				The proportion of tertiary industry in GDP, D10	0.075
Social indicators, B1	0.2	Pollution control, C4	0.102	Harmless treatment rate of domestic garbage, D9	0.034
				Comprehensive utilization rate of industrial solid waste, D8	0.034
				Industrial wastewater discharge compliance rate, D7	0.034
		Resource allocation, C3	0.039	Registered unemployment rate of urban population, D6	0.007
				Housing area per capita, D5	0.032
		Social security, C2	0.045	Education investment as a proportion of GDP, D4	0.034
				Number of hospital beds per 10,000 people, D3	0.011
		Demographic factors, C1	0.013	Urbanization level, D2	0.002
				Natural population growth rate, D1	0.011

TABLE 7: Classification of a comprehensive index of urban ecological environment quality.

Level	Comprehensive index	Comments
1	≥0.8	Strong sustainable development
2	0.65~0.8	Medium sustainable development
3	0.35~0.65	Weak sustainable development
4	0.20~0.35	Sustainable development is hindered
5	≤0.2	Sustainable development is severely hindered

mudslides in 2014, social indicators, economic indicators, and natural indicators all fluctuate to varying degrees, as shown in Figure 1. From the perspective of the comprehensive index of ecological environment quality, it has developed from 0.337 in 2013 to 0.412 in 2018. The overall level is still low, but the development speed is relatively stable. From the perspective of the index classification, the urban ecological environment has been upgraded from level 4 in 2013 to level 3 in 2018. Most of the indicators have developed to a certain extent, indicating that the city's ecological environmental quality is gradually improving.

In the overall environmental development of the city, the overall development of the social environment is relatively stable. The population growth has been effectively controlled, various social infrastructure construction has been gradually improved, the unemployment rate has been drastically reduced, and pollution control has achieved certain results. The overall economic environment is developing slowly. Due to summer floods and mudslides in 2014, the city actively promoted economic system reforms and industrial restructuring which greatly improved people's living standards. The natural environment has developed relatively rapidly. Due to the impact of natural disasters, the natural ecological environment has been severely damaged. However, with the active advancement of postdisaster reconstruction work, the ecological restoration of the natural environment has developed rapidly, especially in the greening construction.

3.2. Analysis of the Comprehensive Index of Social Environment. During the period from 2013 to 2018, the overall social and environmental quality index of the city showed a steady development trend. The social and environmental infrastructure was continuously improved, and the living standards of the people continued to improve. However, the overall development of the social environment was still slow, as shown in Figure 2.

In 2013 and 2018, the city's natural population growth rate remained negative, while the population growth rate in 2018 has risen to 1.2%. The population has increased year by year, but the overall growth rate is controlled within a
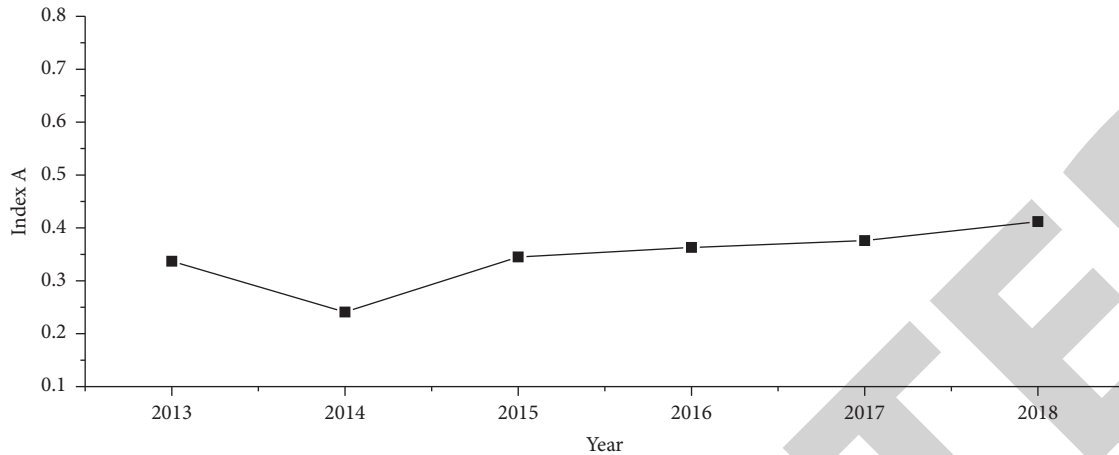
FIGURE 1: Change curve of a comprehensive index of ecological environment quality.
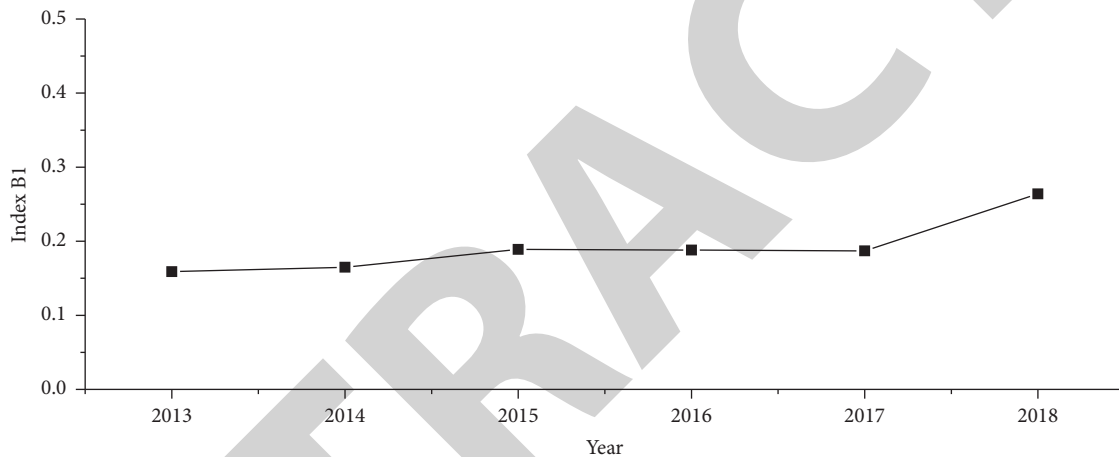
FIGURE 2: The change curve of the comprehensive index of social environmental quality.

reasonable range. The level of urbanization is still lower than the standard value of 70% in moderately developed countries. The population is the main body of the urban ecological environment. Too much denseness will cause pressure on space and the environment [20].

The social public infrastructure is continuously developed and improved, but the overall level still has a certain gap with the standards of the moderately developed countries. The proportion of education investment has increased relatively, but the overall level is still lower than the corresponding standard level [21]. In 2018, the city's per capita housing area reached 36.1 square meters, far higher than the standard value of 16 square meters, and the unemployment rate dropped from 2.4% in 2013 to 0.27% in 2018. The quality of life of the people has improved. In recent years, due to the needs of urban development, the accelerated development of industrial industries has caused certain damage to the urban environment. Environmental protection has received more and more attention. The city has increased its pollution control efforts, vigorously rectified environmental pollution problems, and implemented corresponding measures for related enterprises. Corrective measures have been taken and some results have been achieved [22].

3.3. Analysis of Comprehensive Index of Economic Environment.
From 2013 to 2018, the overall economic and environmental quality index of the city showed a steady and slow development trend. In 2014, due to the impact of natural disasters, the overall economic and environmental level dropped to the level equivalent to two years ago, as shown in Figure 3. The impact of natural disasters on the environment is measured by the following five indicators: (1) direct economic losses caused by geological disasters; (2) investment in geological disaster prevention and control; (3) direct economic loss caused by forest fire; (4) comprehensive control rate of forest diseases, pests, and rodents; (5) direct economic loss caused by environmental pollution and destruction. With the progress of postdisaster reconstruction, the economic environment has gradually recovered and developed, but the overall level has not improved much.

Due to the impact of natural disasters in 2014, the city's economic development was hit hard, the overall level dropped sharply, the city's industrial structure was severely damaged, and the GDP growth rate in the same year showed negative growth. From the analysis of the major natural disasters that have occurred to human beings, it shows that the direct loss of natural disasters does not have a significant
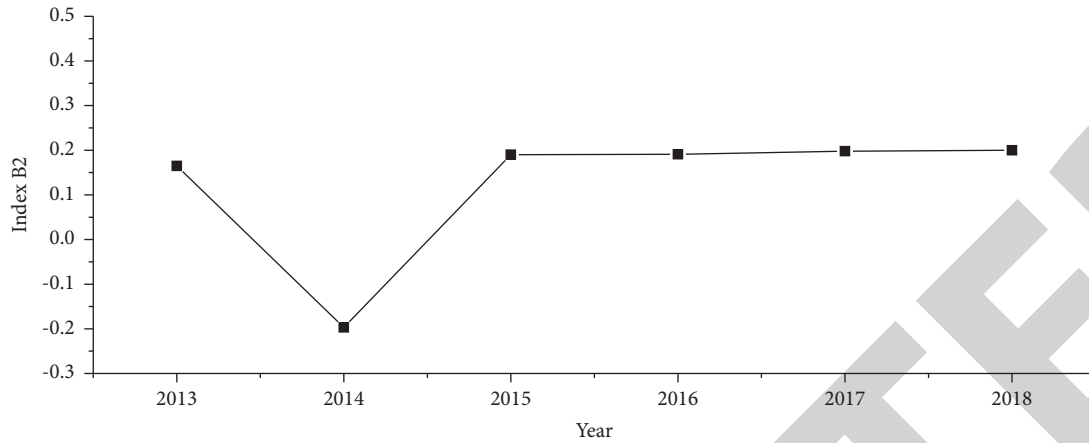
FIGURE 3: The change curve of the comprehensive index of social environmental quality.
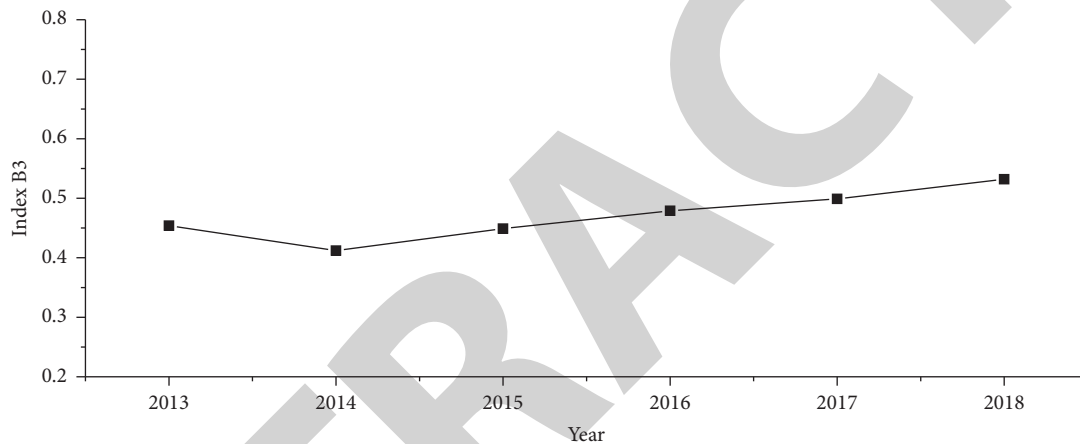
FIGURE 4: The change curve of the comprehensive index of natural environmental quality.

impact on the national economy. General natural disasters have an impact in the short term, and it is difficult to constitute a fundamental disturbance to the huge capital market of the whole country. Since then, until 2018, various economic indicators have gradually recovered, but the development speed has remained slow. Per capita GDP exceeded the standard value for the first time in 2017, and increased significantly in the subsequent period, reaching 48,518 yuan in 2018. In 2017, the annual per capita disposable income reached the relative standard value requirement and exceeded a certain level in 2018. The development trend is good, marking the great improvement and improvement of the quality of life of the people. There is still a certain gap in the urban-rural income ratio, which is related to the low degree of urbanization and the excessive density of the rural population [23].

3.4. Analysis of the Comprehensive Index of Natural Environment.

From 2013 to 2018, the city's natural environmental quality comprehensive index showed an upward trend year by year, as shown in Figure 4. However, due to the impact of summer floods and mudslides in 2014, the urban natural environment was severely damaged, the structure of

the ecosystem was changed, and the function of the ecosystem was greatly affected.

Before the floods and mudslide disasters, the city's socio-economic development was stable, and the trend was good. Financial revenue has increased steadily, the people's income has increased rapidly, and the quality of life has improved rapidly. The industrial economy has developed rapidly, the scale of investment in the city has continued to expand, and the overall level of urban prices has been relatively stable. After the natural disaster, the social economy was severely damaged, especially the secondary industry suffered huge losses. The total agricultural production decreased, and the main crops were affected by the dual impact of floods and mudslides. Large-scale collapses, landslides, and corresponding formation of mud-rock flows and dammed lakes have been caused in the mountainous area, and the natural ecological environment has been severely damaged [24]. Animal habitats, rare animals and plants, native vegetation, and the living environment of wild animals and plants have all been severely affected, the structure of the ecosystem has changed, and the function of the ecosystem has been damaged.

3.5. Urban Development Suggestions. It is necessary to promote urban ecological environment protection through scientific and technological innovation, and vigorously research, develop and promote key applicable technologies for urban environmental protection in the main urban areas. We should fully publicize environmental awareness to the public, give full play to the role of news media and social organizations, strengthen public publicity through various media, and popularize relevant policies and regulations on ecological and environmental protection. And, let the masses imperceptibly contact and accept the correct environmental protection awareness and concepts in their daily life and consciously participate in the actual actions of environmental protection.

4. Conclusion

Cities are gradually developed on the basis of adapting and transforming the natural environment. In a certain urban area, human activities, natural environment, and other factors and their mutual influence constitute the urban ecological environment. Therefore, the evaluation of urban ecological environment quality is of great significance to the analysis of urban development.

(1) This paper chooses to use the analytic hierarchy process to determine the index weights. After calculation, the comprehensive index of the ecological environment quality of a city in western China has developed from 0.337 in 2013 to 0.412 in 2018, indicating that the urban ecological environment of this city has changed from 2013 to 2018. The current ecological environment quality of the city is average, and the economy, society, and environment are not yet in harmony. From the perspective of index classification, the ecological environment of the city has been upgraded from the level 4 standard in 2013 to the level 3 standard in 2018. The overall level is still low, but the urban ecological environment is gradually improving, maintaining a certain degree of sustainability.

(2) During the period from 2013 to 2018, the comprehensive capacity of the ecological environment of a city in western China generally showed a steady increase, except for the impact of floods and mudslides in 2014. The overall level of the comprehensive index of ecological environment quality is still low, but the development speed is relatively stable. The overall development of the social environment is relatively stable, the overall economic environment develops slowly, and the natural environment develops relatively rapidly. Although affected by the natural disasters in 2014, due to the active advancement of postdisaster reconstruction, the ecological restoration of the natural environment has developed rapidly, and the urban environment has been improved. [25, 26].

Data Availability

The figures and tables used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] W. Li, Z. Ouyang, X. Meng, and X. Wang, "Plant species composition in relation to green cover configuration and function of urban parks in Beijing, China," *Ecological Research*, vol. 21, no. 2, pp. 221–237, 2006.

[2] I. Esau, L. Bobyl ev, V. Donchenko et al., "An enhanced integrated approach to knowledgeable high-resolution environmental quality assessment," *Environmental Science & Policy*, vol. 122, no. 9991, pp. 1–13, 2021.

[3] G. Andria, G. Cavone, and A. M. L. Lanzolla, "Modelling study for assessment and forecasting variation of urban air pollution," *Measurement*, vol. 41, no. 3, pp. 222–229, 2008.

[4] J. Liu, M. Ma, F. Zhang, Z. Yang, and J. Domagalski, "The ecohealth assessment and ecological restoration division of urban water system in Beijing," *Ecotoxicology*, vol. 18, no. 6, pp. 759–767, 2009.

[5] B. Fang, C. F. Liu, L. L. Zou, and Y. M. Wei, "The assessment of health impact caused by energy use in urban areas of China: an intake fraction-based analysis," *Natural Hazards*, vol. 62, no. 1, pp. 101–114, 2012.

[6] S. Miethaner, F. Koenig, and B. Lehmann, "A new approach for the assessment of urban waters," *Wasserwirtschaft*, vol. 98, no. 5, pp. 21–26, 2008.

[7] A. M. Rages, R. Oldenkamp, N. L. Preeker, J. Wernicke, and U. Schlink, "Cumulative risk assessment of chemical exposures in urban environments," *Environment International*, vol. 37, no. 5, pp. 872–881, 2011.

[8] A. Ciaramella, V. Puglisi, and T. Truppi, "Environmental performance assessment for urban districts," *Journal of Place Management and Development*, vol. 7, no. 1, pp. 74–89, 2014.

[9] D. dministration and A. U. Development, "Social and environmental assessment management framework," *Vdm Verlag Dr Müller*, vol. 26, no. 4, pp. 41–53, 2004.

[10] M. A. Musse, D. A. Barona, and L. M. S. Rodriguez, "Urban environmental quality assessment using remote sensing and census data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 71, pp. 95–108, 2018.

[11] S. Lai and C. Zoppi, "Un'ontologia dei processi di valutazione ambientale strategica per i piani urbanistici," *Scienze Regionali*, vol. 11, no. 1, pp. 131–138, 2012.

[12] M. Meneses, J. C. Pasqualino, and F. Castells, "Environmental assessment of urban wastewater reuse: treatment alternatives and applications," *Chemosphere*, vol. 81, no. 2, pp. 266–272, 2010.

Research Article

Quantitative Detection of Gastrointestinal Tumor Markers Using a Machine Learning Algorithm and Multicolor Quantum Dot Biosensor

Gaowa Saren [iD],[1] **Linlin Zhu,**[2] **and Yue Han**[3]

[1]*The Department of Medical Nursing, Hulunbuir Vocational Technical College, Hulunbuir 021000, China*
[2]*Department of Pathology, Hulunbeir People's Hospital, Hulunbuir 021008, China*
[3]*Department of Endoscopy, Hulun Buir Weikang Digestive Hospital, Hulunbuir 021000, China*

Correspondence should be addressed to Gaowa Saren; sarengaowa@hlbrzy.com

This work was to explore the application value of gastrointestinal tumor markers based on gene feature selection model of principal component analysis (PCA) algorithm and multicolor quantum dots (QDs) immunobiosensor in the detection of gastrointestinal tumors. Based on the PCA method, the neighborhood rough set algorithm was introduced to improve it, and the tumor gene feature selection model (OPCA) was established to analyze its classification accuracy and accuracy. Four kinds of coupled biosensors were fabricated based on QDs, namely, 525 nm Cd Se/Zn S QDs-carbohydrate antigen 125 (QDs525-CA125 McAb), 605 nm Cd Se/Zn S QDs-cancer antigen 19-9 (QDs605-CA19-9 McAb), 645 nm Cd Se/Zn S QDs-anticancer embryonic antigen (QDs 645-CEA McAb), and 565 nm Cd Se/Zn S QDs-anti-alpha-fetoprotein (QDs565-AFP McAb). The quantum dot-antibody conjugates were identified and quantified by fluorescence spectroscopy and ultraviolet absorption spectroscopy. The results showed that the classification precision of OPCA model in colon tumor and gastric cancer datasets was 99.52% and 99.03%, respectively, and the classification accuracy was 94.86% and 94.2%, respectively, which were significantly higher than those of other algorithms. The fluorescence values of AFP McAb, CEA McAb, CA19-9 McAb, and CA125 McAb reached the maximum when the conjugation concentrations were 25 μg/mL, 20 μg/mL, 30 μg/mL, and 30 μg/m, respectively. The highest recovery rate of AFP was 98.51%, and its fluorescence intensity was 35.78 ± 2.99, which was significantly higher than that of other antigens ($P < 0.001$). In summary, the OPCA model based on PCA algorithm can obtain fewer feature gene sets and improve the accuracy of sample classification. Intelligent immunobiosensors based on machine learning algorithms and QDs have potential application value in gastrointestinal gene feature selection and tumor marker detection, which provides a new idea for clinical diagnosis of gastrointestinal tumors.

1. Introduction

Tumors are caused by abnormal tissue growth caused by genetic mutations and other factors, and malignant tumors are often referred to as cancers. They invade and destroy surrounding tissue, can lead to metastasis and abnormal growth, and, if left untreated, can pose a huge threat to human health. The accurate and reliable classification of tumors is the key to the successful diagnosis and treatment of cancer [1]. In recent years, with the successful application of feature selection in bioinformatics, especially in the face of

many high-dimensional data classification tasks, it has shown ideal performance [2, 3]. However, due to the complexity and variability of gene expression profile datasets and "dimension disaster" and other problems, tumor characteristic gene selection algorithms generally have shortcomings such as high computational complexity and low classification accuracy [4]. Principal component analysis (PCA) is an optimal orthogonal transformation based on the statistical properties of the target. It is the best transformation because it has important good properties. The new components generated after the transformation are

orthogonal or irrelevant, and some new components represent the minimum mean square error of the original vector. After transformation, the vector becomes more determinate, and the energy is more concentrated, which makes it very important in feature extraction and data compression. Some studies have pointed out that PCA simply reduces the spatial dimension in the process of feature gene selection without considering the correlation between features and categories, which is a blind feature selection method [5] and therefore needs to be further optimized.

Gastrointestinal tumors mainly include esophageal cancer, gastric cancer, and colorectal cancer and are one of the common malignant tumors in humans. The global incidence of gastrointestinal tumors accounts for approximately 20% of all tumors [6], of which 35% of patients with malignant tumors die because of intestinal tumors, and the prognosis is poor [7]. In China, the proportion of patients who died of gastrointestinal tumors ranks among the top five in malignant tumors [8]. At present, the diagnosis of gastrointestinal tumors mainly uses imaging methods such as endoscopy, air-barium double contrast, computed tomography (CT), and magnetic resonance imaging (MRI). The pathological diagnosis result is used as the gold standard, but there are certain limitations in the diagnosis of early latent cancer under imaging and pathology [9]. With the continuous research of molecular biology and tumor molecular mechanisms in recent years, tumor marker detection has been used as an effective means of gastrointestinal cancer diagnosis because of its low cost, noninvasiveness, and simple operation [10]. Tumor marker detection can effectively identify benign and malignant diseases and tumor staging and can also detect tumor recurrence and metastasis [11].

Carcinoembryonic antigen (CEA) is a protein complex rich in polysaccharides. Current research results have found that, under normal circumstances, CEA in the body can be metabolized through the gastrointestinal tract. Once a gastrointestinal tumor occurs, CEA metabolism in the body will be abnormal, which will lead to a significant increase in the CEA content in the body [12]. At present, CEA has been used as a tumor marker for a variety of cancers, such as colorectal cancer, gastric cancer, lung cancer, and breast cancer, and it plays an important role in early tumor screening. Carbohydrate antigen 125 (CA125) is a group of high-molecular-weight glycoproteins mainly found in the epithelial cells of the digestive tract, the endothelium of the pleural oviduct, and the endocervical lining [13]. At present, CA125 is mainly used in the diagnosis and screening of ovarian cancer. In recent years, a large number of research results have pointed out that the levels of CA125 in cancer patients, such as those with gastric cancer, colorectal cancer, and breast cancer, have increased to varying degrees [14]. Cancer antigen 19-9 (CA19-9) is a monosialic acid ganglioside. A large number of studies have shown that the levels of CA19-9 in patients with adenocarcinoma, gastric cancer, and colorectal cancer are significantly increased [15]. At present, CA19-9 is mainly used clinically in the auxiliary diagnosis and early screening of pancreatic cancer. Alpha-fetoprotein (AFP) is a glycoprotein whose content is extremely low in normal bodies, but its content is significantly higher in cancerous cells. It is currently mainly used in the auxiliary diagnosis of liver cancer. Studies have pointed out that the levels of AFP in the serum of cancer patients, such as gastric cancer, lung cancer, and colorectal cancer patients, are also significantly increased [16]. Positive rates of tumor markers CEA, CA125, CA19-9, and AFP in gastrointestinal tumors have been reported. Studies have pointed out that the positive rate of CEA in gastrointestinal cancer serum is approximately 40% [17], the detection rate of CA125 is approximately 15% [18], the positive rate of CA19-9 is approximately 40% [19], and the positive rate of AFP is approximately 10% [20]. The detection rates of different tumor markers are different, and the sensitivity and specificity of a single tumor marker are not high. Therefore, the combined determination of multiple markers should be used to increase the detection rate of tumors.

At present, the clinical detection methods of tumor markers mainly include ELISA, chemiluminescence immunoassays, biochip methods, and microarray methods [21]. However, these detection methods are costly, time-consuming, and labor-intensive. Lateral flow immunoassay (LFIA) is a biosensor made from materials such as cellulose or nitrocellulose. Because of its good stability, high detection sensitivity, and fast detection speed, it is widely used in many fields, such as food safety, water quality detection, and medical diagnosis [22, 23]. Quantum dots (Ds) are semiconductor nanoparticles with a particle size between 1 and 10 nm [24]. QDs have good biocompatibility and luminescence properties and are widely used in the detection of microorganisms, proteins, and nucleic acids [25]. Researchers have used QDs in immunochromatographic techniques and applied them in drug, environmental, food, and medical testing, and they found that the sensitivity of QDs to be tested is significantly improved [26]. Multiple QDs can be excited at the same time, and the emission spectra do not easily overlap [27], which lays a theoretical foundation for the combined determination of multiple markers.

In summary, PCA algorithm has significant advantages in the process of data feature extraction, but it still has limitations in gene feature extraction, which needs to be further optimized. Therefore, based on PCA calculation, the gene feature selection model was established by optimizing it, and the classification accuracy and accuracy of OPCA model based on PCA algorithm in gene feature selection were discussed, so as to provide reference for gene feature selection method selection. In addition, the combined detection of multiple markers is of great significance for tumor monitoring, and QDs have the advantage of being harnessed in immunochromatography. In this study, based on the needs of tumor detection and the advantages of QDs, a multicolor QD-labeled biofilm sensor was developed to realize the quantitative detection of multiple tumor markers and was applied to the detection of GITMs to explore the application value of QD-based multicolor biosensing films in the detection of GITMs and provide a new diagnostic method for the monitoring of gastrointestinal tumors.

2. Experimental Details

2.1. Establishment of a Tumor Gene Feature Selection Model Based on Principal Component Analysis (PCA). A_1, A_2, ..., A_p are set as p-dimensional random variables, and the covariance matrix is expressed as follows:

$$\sum = \left(\sigma_{ij}\right)_{pxp} = E\left[(A - E(X))(A - E(X))^T\right]. \quad (1)$$

A new variable B can be obtained by orthogonal transformation of variable A, and its expression is as follows:

$$Y_p = l_p^T A = l_{p1}A_1 + l_{p2}A_2 + \cdots + l_{pp}A_p. \quad (2)$$

The principal component analysis covariance matrix is expressed as follows:

$$Cov\left(Y_i, Y_j\right) = Cov\left(l_i^T A, l_j^T A\right) = l_i^T A_j^l, j = 1, 2, \cdots, p, \quad (3)$$

where l_i represents P constant vectors, $i = 1, 2, \cdots, p$.

The PCA feature selection method has the disadvantage of weak gene discrimination ability in the process of feature gene selection, and its gene microarray contains many redundancies. To overcome the above shortcomings, a neighborhood rough set algorithm was introduced to optimize it, and a tumor gene feature selection model was established, which was named OPCA.

A neighborhood rough set is an iterative process in neighborhood calculation with high computational time complexity [28]. It is assumed that a gene dataset contains Z samples and m genes, and the reduction time consumption is P. To reduce the iterative operation, the OPCA model used principal component analysis dimensionality reduction to construct the low-dimensional feature space and then used a multineighborhood rough algorithm to screen out the better feature gene set. For a given n-dimensional feature space Q, its output covariance matrix H is expressed as shown in (3), and M represents the sample size of the input space.

$$H = \frac{1}{M} \cdot \sum_{k=1}^{M} x_k x_k^T. \quad (4)$$

If D is the N-order matrix and β is a real number, then there is the following equation:

$$DX = \beta X, \quad (5)$$

where β is called the eigenvalue of D and X is the eigenvector of D.

For a given gene dataset G, the radius parameter and the lower limit parameter of its attribute domain were calculated. The PCA algorithm was used to calculate the contribution rate of the gene dataset, and the gene dataset with a contribution rate greater than 1% was selected to initialize the reduced dataset and calculate the attribute domain value. The feature set was used to represent the gene attribute column in the gene dataset whose contribution rate was more than 1%. Then, the positive domain of the attribute and its importance degree were calculated, the positive domain set of the attribute was obtained, and the importance degree was judged to be greater than the lower limit parameter.

According to the judgment result, the better characteristic gene set was finally output. The characteristic gene selection process of OPCA based on PCA is shown in Figure 1.

2.2. Validation of the Tumor Gene Feature Selection Model Based on PCA. Colon tumor (https://featureselection.asu.edu/datasets.php) and gastric cancer microarray dataset GSE54129 were selected to verify the OPCA gene feature selection model. There were 2,000 colon tumor dataset features. The positive and negative sample sizes were 22 and 40, respectively, and the gastric cancer microarray dataset GSE54129 contained 3894 features and 21 normal and 111 abnormal sample sizes, respectively. The experimental environment was CPU: AMD Athlon™ II X4 645 processor; memory: 4 G; system: Windows 7; experimental software: MATLAB-R2010A and Weka-3.9.0. The OPCA model neighborhood parameters ranged from 0 to 2, and the importance lower limit was 0.01. The number of characteristic genes in both the colon tumor and gastric cancer datasets was 6, and the thresholds were 0.77 and 0.19, respectively. Under the same conditions, the number of characteristic genes in the OPCA gene feature selection algorithm established in this study was compared with the improved genetic algorithm (IGA) [29], PCA [30], and neighborhood rough set (NRS) [2]. The classification precision and classification accuracy were compared and analyzed.

2.3. Selection of Multicolor QDs. After chemical modification, water-soluble QDs can be covalently combined with antibodies (Ab) to form stable QD-Ab conjugates. One microliter of 525 nm, 545 nm, 565 nm, 585 nm, 605 nm, 625 nm, and 645 nm Cd Se/ZnS QDs (Ocean Nanotech, USA) was dropped onto a nitrocellulose (NC) membrane. After drying, the emission spectrum was measured at 365 nm using a spectrophotometer (Shimadzu, Japan).

2.4. Preparation of QD-Ab Antibody Conjugate. After 35 μL of 525 nm, 565 nm, 605 nm, and 645 nm Cd Se/ZnS QDs with a concentration of 8 mol/L was added to different centrifuge tubes, certain amounts of 1-(3-dimethylamino-propyl)-3-ethylcarbodiimide hydrochloride 1-(3-(ethyl-iminomethylideneamino)-N, N-dimethylpropan-1-amine, hydrochloride, EDC) (American Sigma-Aldrich company) and N-hydroxysulfosuccinyl Imine (N-Hydroxysulfosuccinimide, NHS) were added to mix well. Phosphate buffer with a concentration of 0.01 mol/L was added to make the final concentrations of EDC and NHS in solutions of 0.4 mg/mL and 0.2 mg/mL, respectively, and shaken for 30 min at room temperature for activation. AFP McAb was added to the activated 565 nm Cd Se/ZnS QDs, CEA McAb was added to the activated 645 nm Cd Se/ZnS QDs, and the reaction was carried out at room temperature for 2 hours. After adding 500 μL of 0.5% bull serum albumin (BSA) to react at room temperature for 1 hour, the mixture was centrifuged at 12,000 rpm/min for 30 minutes to collect the precipitate, which was washed with 0.01 M phosphate buffer
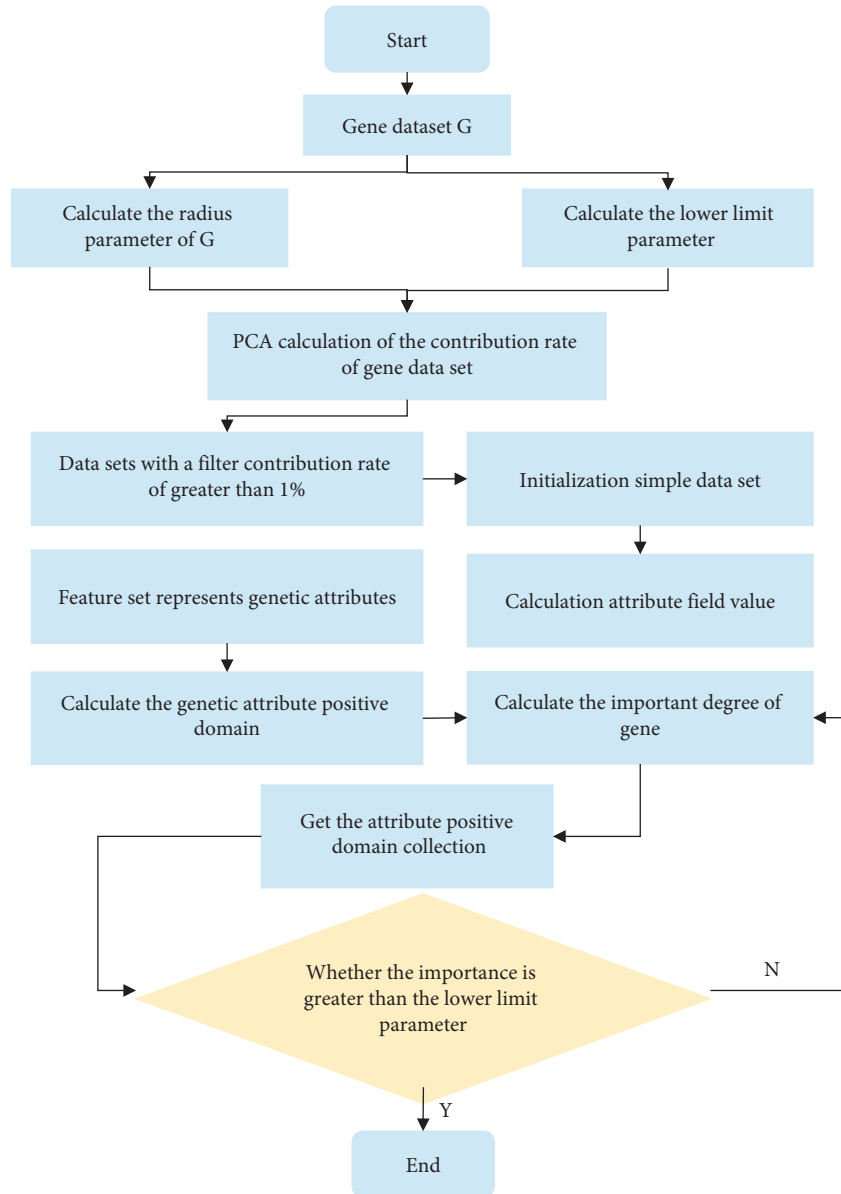
FIGURE 1: Flow chart of characteristic gene selection based on OPCA.

3 times. The pellet was resuspended in 0.01 M phosphate buffer containing 1% BSA, 0.01% NaN₃, and 0.02% Tween-20 to obtain the QD-Ab antibody conjugate. The specific preparation process of the QD-Ab antibody conjugate is shown in Figure 2.

2.5. Exploration and Identification of Optimal Preparation Conditions for QD-Ab Antibody Conjugates.

After activation, $10\,\mu L$ each of 565 nm and 645 nm Cd Se/ZnS QDs was added to a certain amount of phosphate buffer to adjust the pH to 5.0, 5.5, 6.0, 6.5, 7.0, 7.4, 8.0, 8.5, and 9.0. Then, $20\,\mu L$ of AFP McAb and CEA McAb at a concentration of 2 mg/mL was added under different pH systems to react for 2 h at room temperature. The fluorescence intensity of the QD-Ab antibody conjugate prepared under different pH conditions was measured by a fluorescence spectrophotometer.
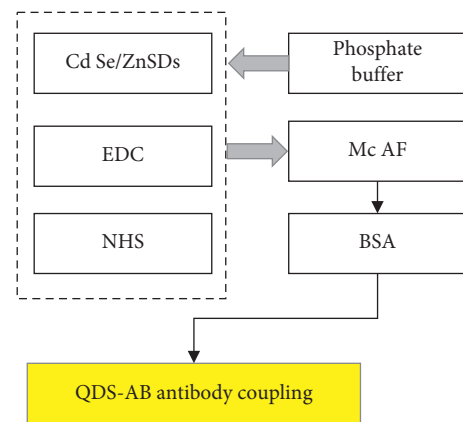
FIGURE 2: Flow chart of the preparation of the QD-Ab antibody conjugate.

After activation, 10 μL of 565 nm and 645 nm Cd Se/ZnS QDs was taken and added to 0, 5, 10, 15, 20, 25, 30, and 35 g of AFP McAb and CEA McAb to adjust the pH of the reaction system to the optimal conditions and then react for 2 hours at room temperature. A fluorescence spectrophotometer was adopted to detect the fluorescence spectrum of each system.

After activation, 565 nm and 645 nm Cd Se/ZnS QDs were mixed with the optimal amounts of AFP McAb and CEA McAb, respectively. At the optimal pH, the solution was allowed to react at room temperature for 30 min, 60 min, 90 min, 120 min, 150 min, and 180 min. The reaction products were collected in a fluorescence spectrophotometer to detect the fluorescence intensity of the QD-Ab antibody conjugate under different reaction times.

Fluorescence spectroscopy and UV absorption spectroscopy were used to analyze the changes in light absorption before and after the coupling of Cd Se/ZnS QDs to identify whether the QD-Ab antibody conjugate was successfully prepared.

2.6. Preparation of QD Biosensors. The glass fiber membrane was used as the binding pad, cut into a size of 0.4 cm × 0.7 cm, and immersed in 0.01 M phosphate buffer containing 2% BSA, 1% sucrose, and 0.1% Tween-20 (pH 7.4). After it was dried at 37°C, 10 μL QDs-Ab-AFP McAb coupling objects, QDs-Ab-CEA McAb coupling objects, QDs-Ab-CA125 McAb coupling objects, and QDs-Ab-CA19-9 McAb coupling objects were added dropwise and dried at 37°C in vacuum. The NC membrane was cut into a size of 0.4 cm × 2.4 cm, and the AFP McAb was diluted with 0.01 M phosphate buffer (pH 7.4) containing 1% sucrose to a final concentration of 1 mg/mL. The 2 μL/cm was determined as the T band, and 0.5 mg/mL goat anti-mouse IgG was used as band C and dried at 4°C for use. The sample pad, bonding pad, NC membrane, and absorbent board were laminated in sequence and pasted onto a polyvinyl chloride (PVC) base plate. After drying at 4°C, four different quantum dot biosensors were obtained.

2.7. Quantitative Detection and Verification of the QD Biosensor. AFP and CEA were diluted with 0.01 M phosphate buffered saline (PBS) (pH = 7.4) to 0.25, 0.5, 1, 5, 10, 40, 60, 80, 100, and 120 ng/mL, respectively. The prepared QD biosensor was used to detect the fluorescence values of AFP and CEA and plotted as a curve. At the same time, the fluorescence of the QD biosensor at different concentrations was observed under a microscope in the dark.

AFP was diluted with 0.01 M PBS (pH = 7.4) to final concentrations of 70 ng/mL, 30 ng/mL, and 3 ng/mL. After 80 μL of diluted AFP, the prepared QD biosensor was adopted to detect the fluorescence value, and the recovery rate of sample detection was calculated according to the standard curve.

Based on the methods introduced by Shariatifar et al. [31], cross-reaction experiments were performed with BSA, CA125, CA-19-9, CEA, and AFP to evaluate whether the test results between AFP and BSA, CA125, CA-19-9, and CEA could affect each other.

2.8. QDs Biosensor to Detect CEA, CA125, CA19-9, and AFP. For simultaneous detection of multiple tumor markers, QDs-Ab-CEA McAb, QDs-Ab-CA125 McAb, QDs-Ab-CA19-9 McAb, and QDs-Ab-AFP McAb should be prepared according to the abovementioned QDs-Ab antibody conjugates. Then, the corresponding antibody was added during the coating process of the detection tape and the quality control tape. For other operation steps, refer to the above steps.

2.9. Statistical Methods. The test data were processed using SPSS 19.0 statistical software, and the data were analyzed by one-way analysis of variance. The measurement data were expressed as the mean ± standard deviation ($\overline{x} \pm s$), and the count data were expressed as a percentage (%), using the χ^2 test. *P < 0.05* indicated that the difference was statistically significant.

3. Experimental Results and Analysis

3.1. Comparison of Tumor Gene Feature Selection Results with Different Gene Selection Algorithms. Under the same conditions, the number of feature genes, classification precision, and classification accuracy of the OPCA gene feature selection algorithm were compared with those of the IGA algorithm, PCA algorithm, and neighborhood rough set (NRS) algorithm. In Figure 3, the number of selected characteristic genes by the OPCA gene feature selection algorithm in both the colon tumor and gastric carcinoma datasets was 6, which was significantly lower than that of the other algorithms. The classification precision of OPCA in the colon tumor and gastric carcinoma datasets was 99.52% and 99.03%, and the classification accuracy was 94.86% and 94.2%, respectively. The OPCA gene feature selection algorithm selected the fewest feature genes, and its classification precision and accuracy were higher than those of the current algorithm. Therefore, the dataset of the original gene expression profile contained much redundant information. In the process of feature gene selection, the OPCA gene feature selection algorithm effectively removed redundant noise, improved the classification ability of the feature gene subset, and extracted fewer feature gene subsets. The reason was that PCA reduced the dimension of the feature space and the complexity of the neighborhood calculation. Meanwhile, the multineighborhood rough set algorithm calculated the neighborhood values of each gene through Euclidean distance, constructed neighborhood sets to calculate the approximation, and extracted the subset of feature genes based on heuristic search. The performance of gene feature extraction was improved by combining the two algorithms. The results showed that the OPCA gene feature selection algorithm was feasible and effective in the process of tumor gene feature extraction.

(a)

(b)

(c)
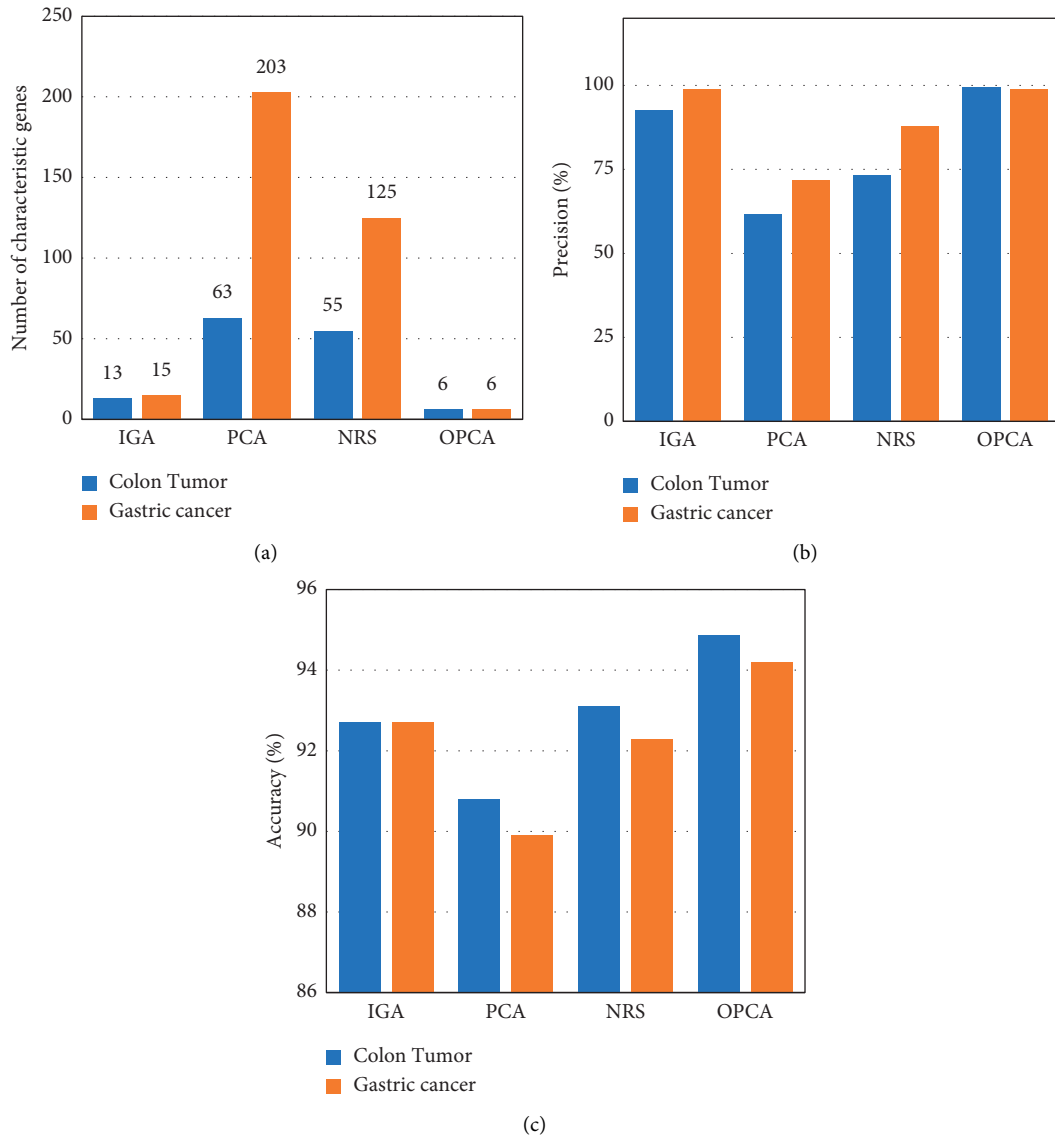
FIGURE 3: Comparison of tumor gene feature selection results with different gene selection algorithms. (a) Comparison of the number of tumor gene characteristics with different gene selection algorithms; (b) comparison of precision of tumor gene feature selection by different gene selection algorithms; (c) comparison of accuracy of tumor gene feature selection by different gene selection algorithms.

3.2. The Fluorescence Spectrum Detection Results of QDs. Water-soluble QDs have univariate excitation, and multiple emission characteristics are the basis for the analysis of multiple markers [32]. The selection principle of multicolor QDs is that the emission spectra of multiple QDs cannot overlap [33]. Spectral analysis of QDs in the range of 525~645 nm (Figure 4) showed that there was obvious spectral overlap between QDs in the adjacent wavelength range. Therefore, wavelengths with a wavelength interval of 40 nm were selected for the experiment in this study; that is, 525 nm, 565 nm, 605 nm, and 645 nm QDs were selected as fluorescent materials for subsequent experiments. After excitation by 365 nm excitation light, nonoverlapping 525 nm, 565 nm, 605 nm, and 645 nm QDs were obtained, indicating that four tumor markers can be quantitatively detected simultaneously, which improved detection efficiency and reduced detection costs.

3.3. Coupling Conditions Analysis of QDs and Antibody. When QDs are adopted to label antibodies, the optimal conditions of the reaction system must be analyzed to ensure the effective activity of the antibodies [34]. The fluorescence intensity of QDs and antibody coupling were analyzed under different pH conditions (Figure 5). As the pH value increased, the fluorescence intensity of QDs coupled with antibodies first increased and then decreased, and the optimal pH values for coupling of different QDs with antibodies had certain differences. The optimal coupling pH of AFP McAb was 7.5, the optimal coupling pH values of CEA McAb and CA19-9 McAb were both 7.0, and the optimal coupling pH of CA125 McAb was 6.5.

The influence of different antibody concentrations on the fluorescence intensity of the conjugate was analyzed, and the results are illustrated in Figure 6. As the antibody concentration in the reaction system increased, the fluorescence
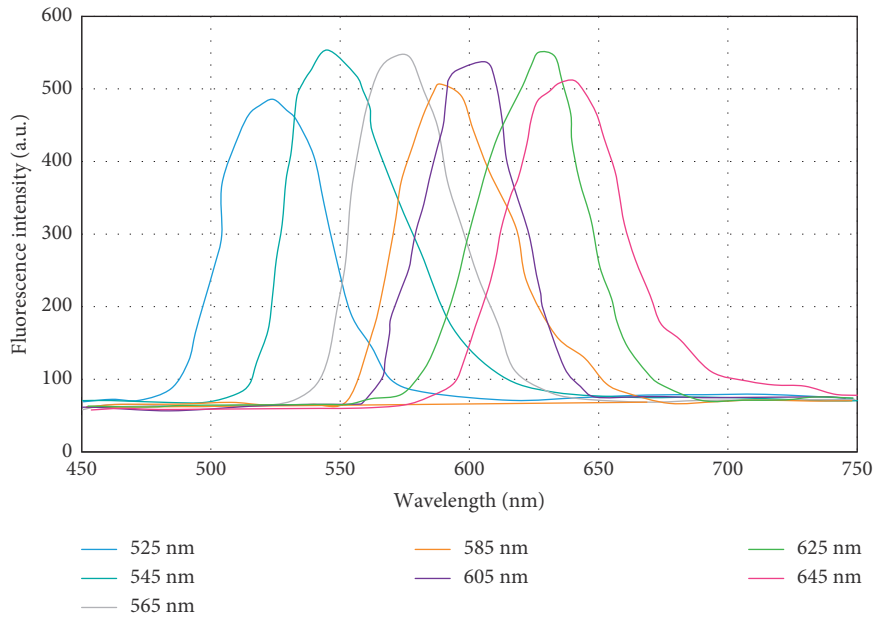
Figure 4: Emission spectra of QDs at different wavelengths.
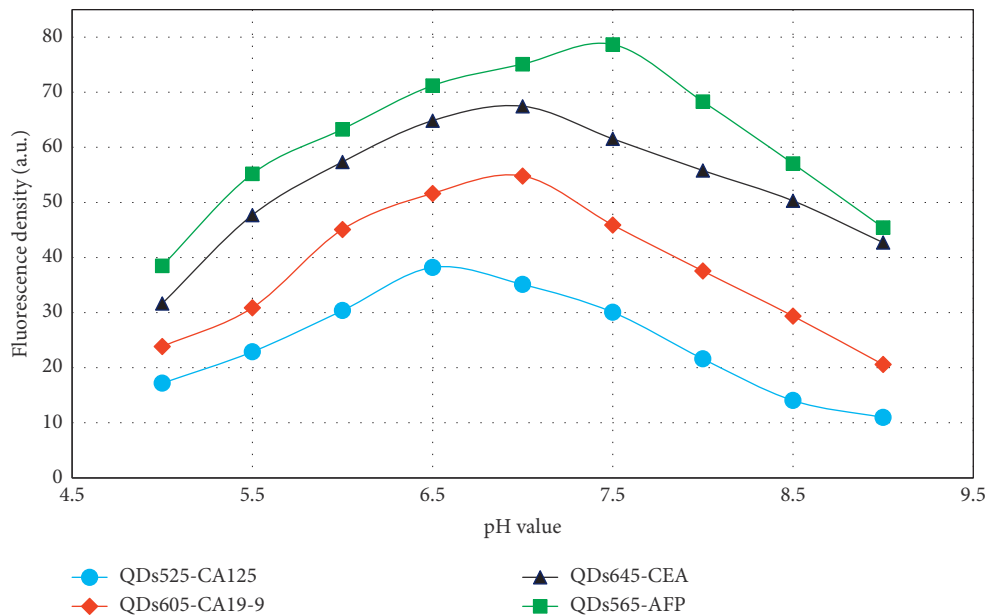
Figure 5: Fluorescence intensity change curve of the conjugate under different pH values.

intensity of the conjugate first increased and then stabilized. The fluorescence values of AFP McAb, CEA McAb, CA19-9 McAb, and CA125 McAb reached maximum values when the coupling concentrations were 25 μg/mL, 20 μg/mL, 30 μg/mL, and 30 μg/mL, respectively. Even if the antibody concentration is increased beyond the maximum value, the fluorescence value of the conjugate will no longer increase. The amount of antibody added has a significant impact on the performance of the conjugate and the entire biosensor [35]. If the amount of antibody in the reaction system is insufficient, it will cause a large number of unreacted sites in the conjugate, and nonspecific binding is likely to occur

during the sample detection process, which will eventually lead to a false positive test result. If the amount of antibody in the reaction system is too high, the remaining antibodies that are not bound by the conjugate will reduce the sensitivity of sample detection when testing the sample [36].

The fluorescence intensity of the coupling substance was detected under different reaction times, as shown in Figure 7. Different antibodies reacted with the coupling substance, and the fluorescence intensity of the antibody coupling substance first increased and then stabilized as the reaction time increased. When the reaction time of AFP McAb and CEA McAb was 90 min, the fluorescence
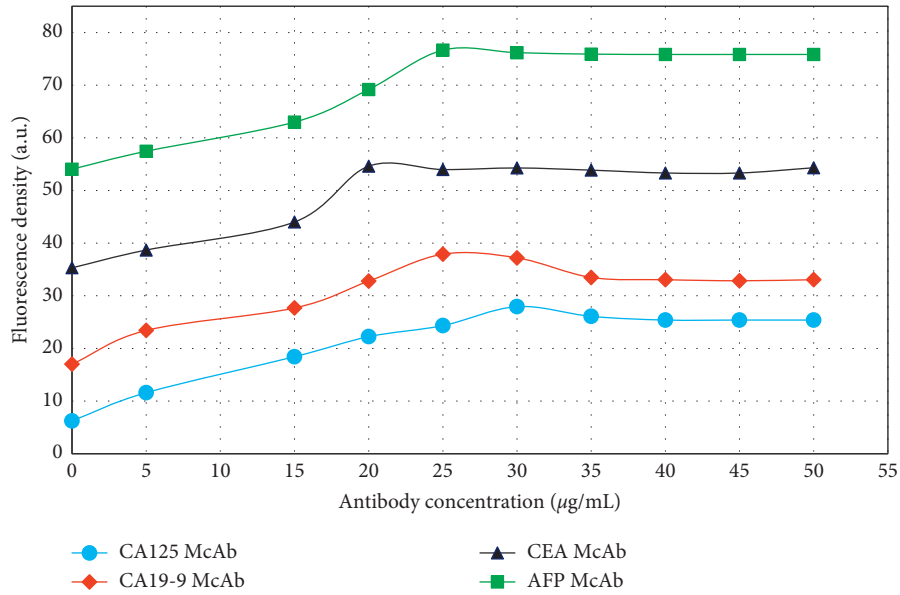
FIGURE 6: The effect of antibody concentration on the fluorescence intensity of the coupling reaction system.
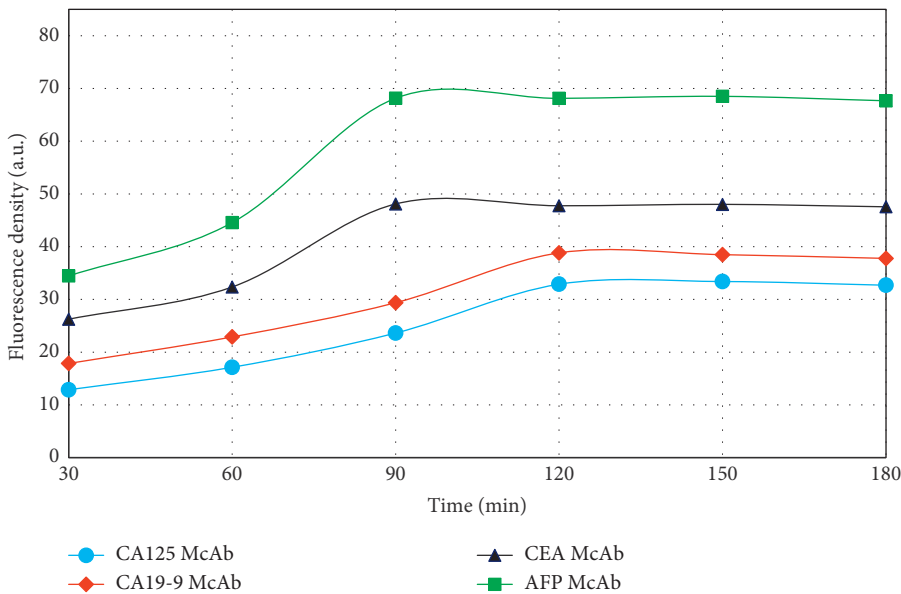
FIGURE 7: The effect of reaction time on the fluorescence intensity of the coupling reaction system.

intensity reached the maximum, and when the reaction time of CA19-9 McAb and CA125 McAb was 120 min, the fluorescence intensity reached the maximum. The maximum fluorescence intensity indicates that the reaction has reached saturation at this time. The reaction time has a significant effect on the reaction of the antibody conjugate. If the reaction time is too short, the reaction between QDs and the antibody conjugate will be insufficient, resulting in false positive results [37]; if the reaction time is too long, the coupling efficiency will decrease.

3.4. Identification of QD-Antibody Conjugates. The successful combination of QDs and antibody conjugates plays an important role in the performance of QD biosensing membranes [38]. The fluorescence spectra of different antibody conjugates were analyzed before and after coupling with QDs, and the results are given in Figure 8. The fluorescence intensities of QDs525-CA125, QDs605-CA19-9, QDs645-CEA, and QDs565-AFP after binding antibodies were more enhanced than those of unconjugated QDs. It may be that the surface defect of QDs is modified after the antibody conjugate is combined with QDs, which increases the prefluorescence intensity of the QD conjugate. Moreover, there is no fluorescence peak shift after the QDs at different wavelengths are combined with the antibody conjugate [39], indicating that the QDs only bind to the antibody conjugate during the coupling process, and there is no polymerization reaction between the QDs.
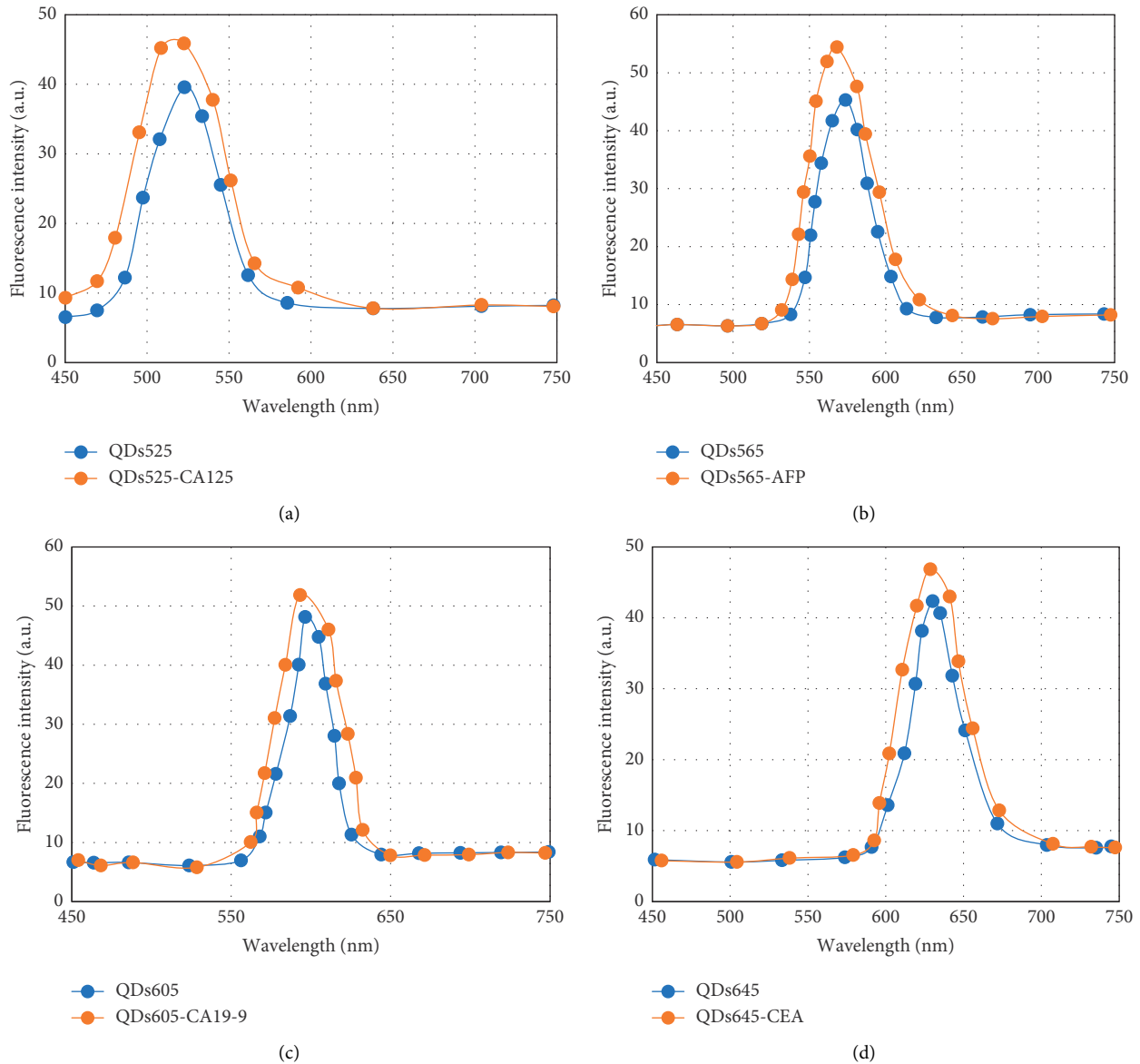
FIGURE 8: Fluorescence spectrum of QDs. (a) Fluorescence spectra of QDs525 and antibody conjugate before and after reaction; (b) fluorescence spectra of QDs565 and antibody conjugate before and after reaction; (c) fluorescence spectra of QDs605-CA19-9 and antibody conjugate before and after reaction; (d) fluorescence spectra before and after the reaction between QDs645-CEA and antibody conjugate.

Figure 9 shows the UV absorption spectra of different antibody conjugates before and after coupling with QDs. The UV absorption values of QDs525-CA125, QDs605-CA19-9, QDs645-CEA, and QDs565-AFP after binding to the antibody were all higher than those of unconjugated QDs, and the maximum absorption wavelength did not shift, indicating that the antibody conjugate and QDs were combined successfully.

3.5. Analysis of AFP Quantitative Detection Results. Under the optimal reaction conditions, the method established in this study was used to detect different concentrations of AFP standards. During the detection of AFP standards, the fluorescence intensity in the fluorescence image increased with increasing AFP concentration (Figure 10). The QD biosensing membrane was placed in a fluorescence reader to detect the fluorescence value of the T-band and C-band, and the AFP standard curve for multicolor QD biosensor detection was analyzed (Figure 11). With increasing AFP concentration, the T/C value of the QD biosensing film increased. The linear fitting equation of the T/C value and AFP concentration was $y = 0.448x - 6.1923$ ($R^2 = 0.9335$), showing a good linear relationship.

Biosensors used in clinical applications need to have repeatable and reliable test results [40]. The actual concentration of AFP, multicolor QD biosensor detection concentration, and recovery rate were further analyzed, as illustrated in Figure 10. The actual concentration of AFP was
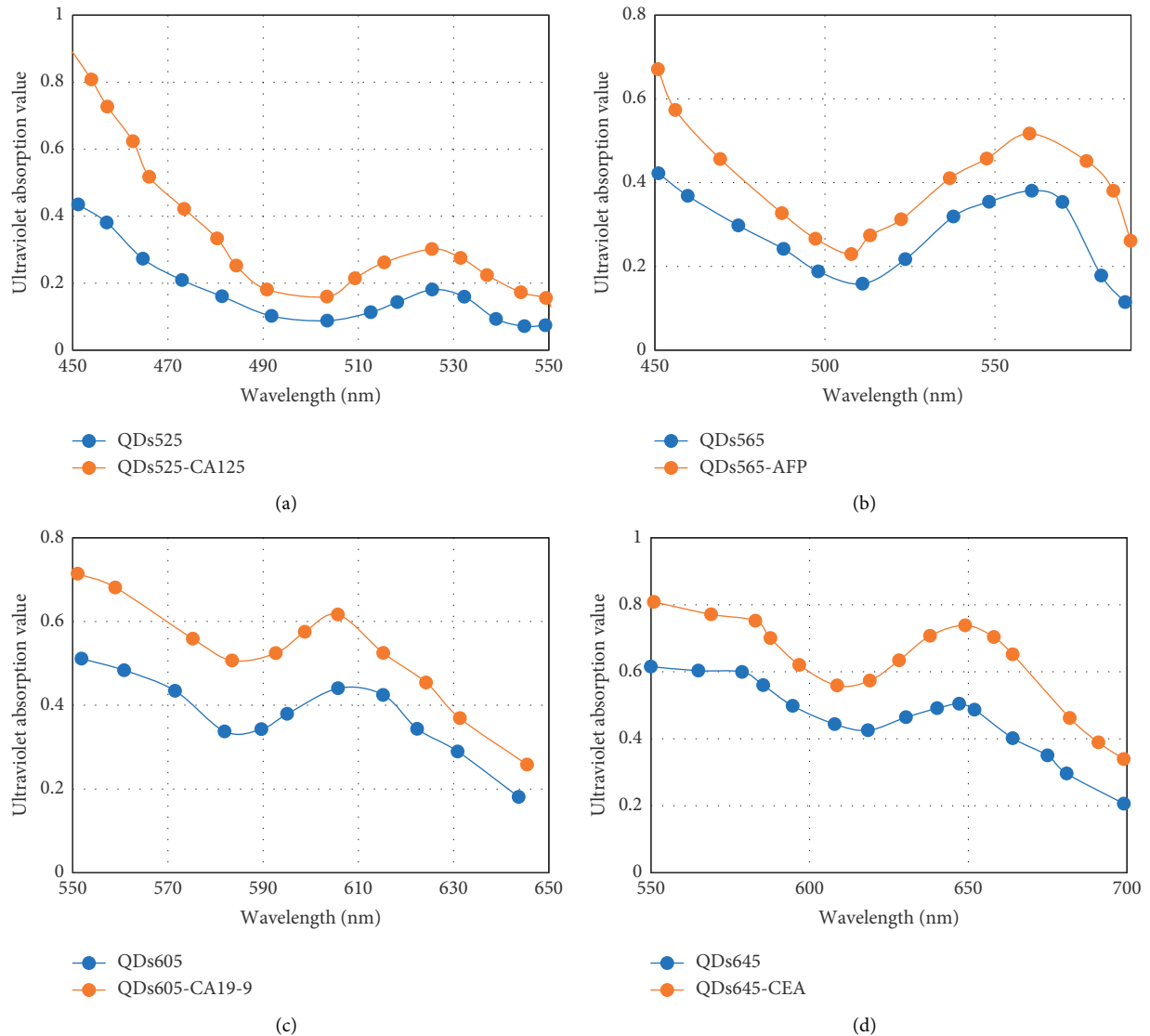
FIGURE 9: UV absorption spectrum of QDs. (a) UV absorption spectrum before and after the reaction of QDs525 with antibody conjugate; (b) UV absorption spectrum before and after reaction of QDs565 with antibody conjugate; (c) UV absorption spectrum before and after reaction of QDs605-CA19-9 with antibody conjugate; (d) UV absorption spectra before and after the reaction between QDs645-CEA and the antibody conjugate.

not much different from the concentration of AFP detected by the multicolor QD biosensor. The recovery rate of AFP was up to 98.51%. The closer the sample recovery rate is to 100% and the closer the detection concentration is to the actual concentration, the higher the reliability of the biosensor is [41], indicating that the multicolor QD biosensor has higher repeatability and reliability.

To verify the specificity of the immune method of the QD biosensor for AFP detection, the QD biosensor was adopted to detect and analyze BSA, CA125, CA-19-9, CEA, and AFP nonspecific antigens (Figure 12(b)). The fluorescence intensity of AFP was 35.78 ± 2.99, which was significantly higher than those of the other antigens ($P < 0.001$). Without labeled AFP, even if the concentration of nonspecific antigen is high, a strong fluorescence value cannot be detected. This shows that when there is no labeled antigen, nonspecific

antigens will not have an immune response, indicating that this method has significant specificity for detecting AFP.

3.6. The Results of Quantitative Detection of GITM Using the Multicolor QD Biosensor.

Cross-reactivity is the main parameter for the specific evaluation of immunoassay methods [42]. For the detection of multiple tumor markers, it is necessary to analyze whether there is a cross-reaction between multiple tumor markers and the biosensor in the same space [33]. In this study, the results of quantitative detection of GITMs using a multicolor QD biosensor were analyzed, and the results are shown in Figure 13. Four independent proteins CA125, CA-19-9, CEA, and AFP were added to the multicolor QD biosensor, and the multicolor QD biosensor diluent was used as a blank control. Then, the detection of
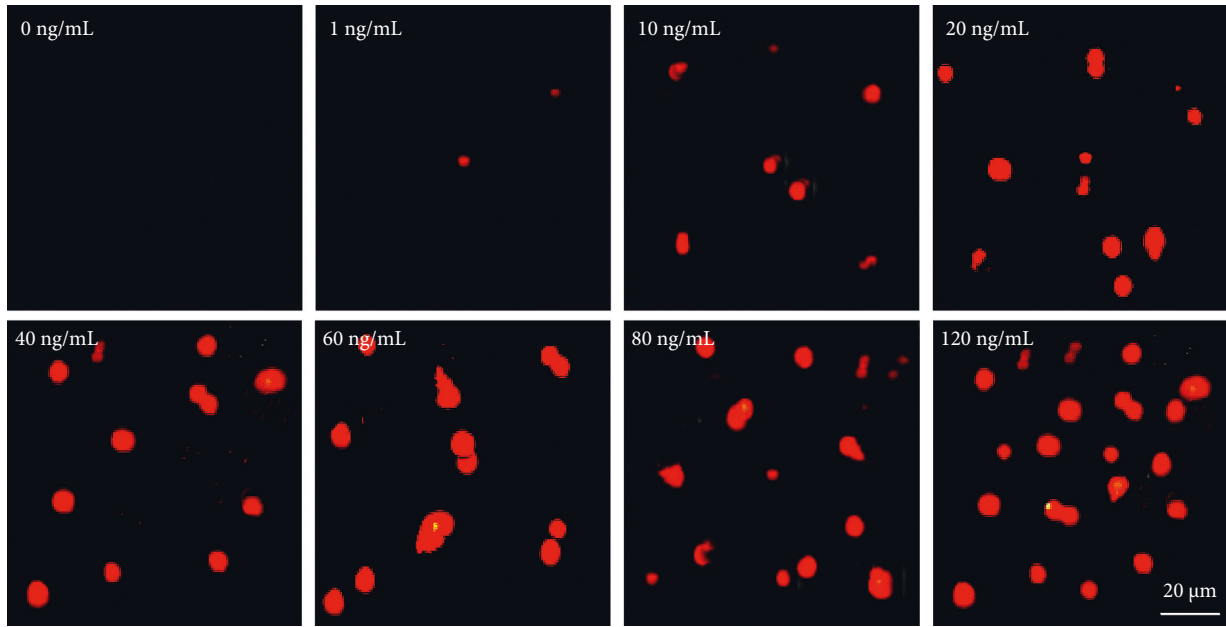
FIGURE 10: AFP fluorescence image detected by QDs sensor.
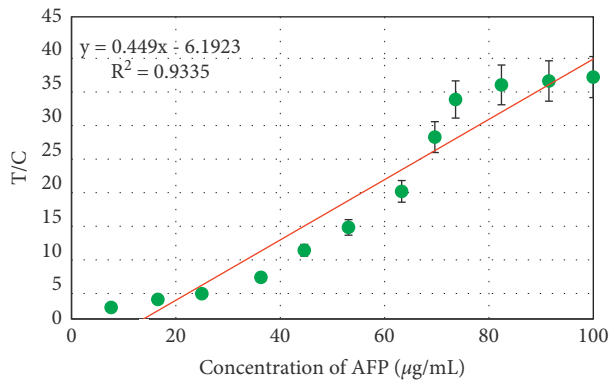
FIGURE 11: AFP standard curve based on the multicolor QD biosensor.
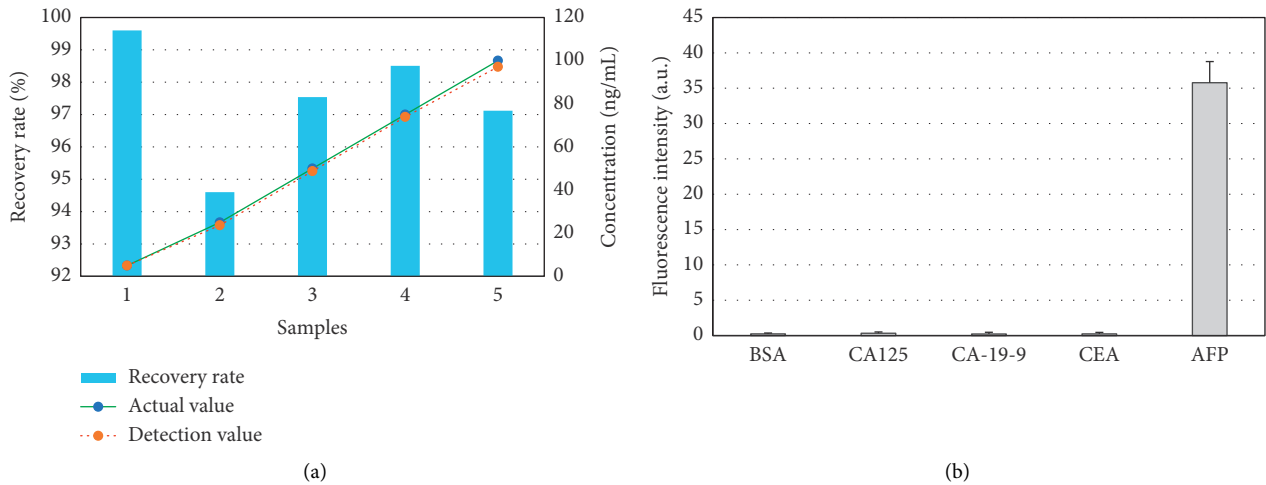
FIGURE 12: Analysis of AFP standard test results. (a) Recovery test result of AFP standard product; (b) detection specificity of AFP.
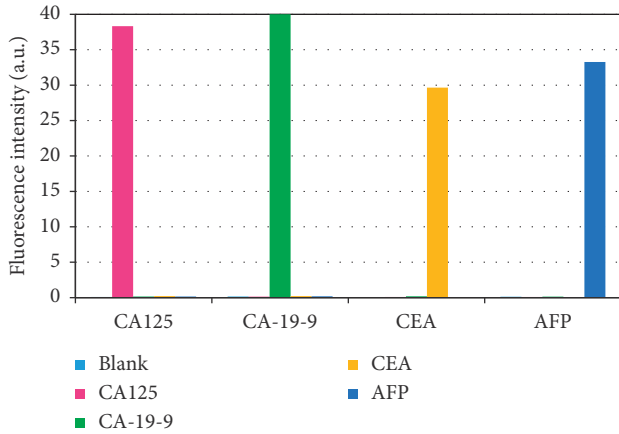
FIGURE 13: Quantitative detection of GITM-specific analysis using a multicolor QD biosensor.
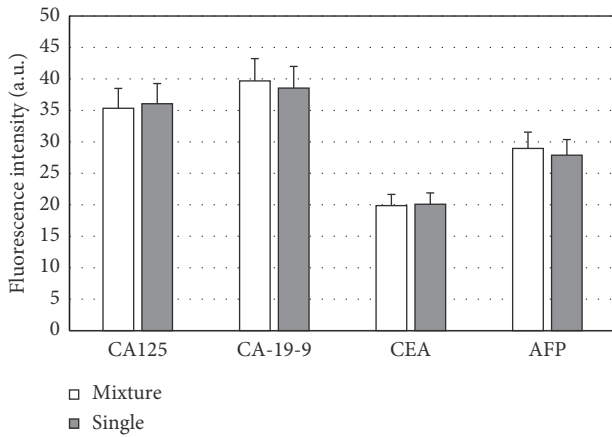
FIGURE 14: Analysis of the results of single index detection and multi-indicator simultaneous detection.
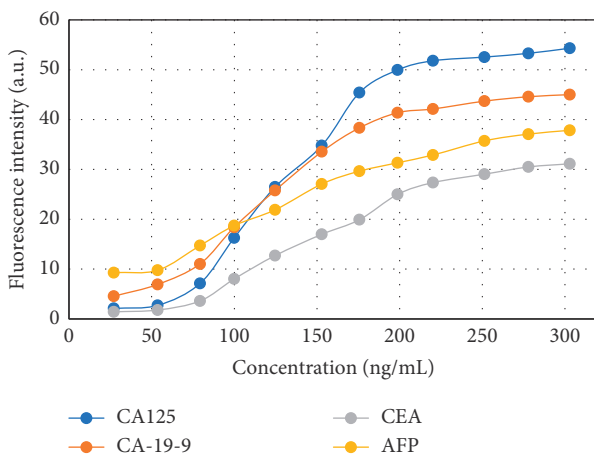
FIGURE 15: The fluorescence spectra of CA125, CA-19-9, CEA, and AFP simultaneously detected by the multicolor QD biosensor.

the four tumor markers showed good specificity, and there was no interference between them.

The results of single index detection and simultaneous multi-indicator detection were compared and analyzed, and

TABLE 1: Linear area table of simultaneous multi-indicator detection using the multiple QD biosensor.

Tumor markers	Standard curve line	Linear working area (ng/mL)	R^2
CA125	$y = 0.2248x - 3.7298$	2.0~51.5	0.9884
CA-19-9	$y = 0.1658x + 2.4854$	4.5~40.0	0.9756
CEA	$y = 0.1262x - 3.3665$	2.0~29.5	0.9637
AFP	$y = 0.1135x + 6.9125$	8.5~36.5	0.9601

the results are shown in Figure 14. Under different methods, there was no significant difference in the fluorescence intensity of the four GITMs ($P > 0.05$), indicating that the specificity of the mixed detection of the four test samples was better.

The fluorescence intensity of the multicolor QD biosensor simultaneously detecting the tumor markers CA125, CA-19-9, CEA, and AFP was analyzed at different concentrations (Figure 15). With increasing concentrations of the tumor markers CA125, CA-19-9, CEA, and AFP, the fluorescence value of the multicolor QD biosensor also showed an upward trend. This is because as the concentration of tumor markers CA125, CA-19-9, CEA, and AFP increases, more immune complexes Ds525-CA125, QDs605-CA19-9, QDs645-CEA, and QDs565-AFP are formed, and the more multicolor QD biosensor T had a higher fluorescence value.

The detection curve was drawn with the concentration of GITMs CA125, CA-19-9, CEA, and AFP as the abscissa and the corresponding T/C value as the ordinate. The linear equation and the corresponding detection linear region are shown in Table 1. When the four tumor markers CA125, CA-19-9, CEA, and AFP were detected by the multiquantum dot biosensor multi-index synchronous detection method, the corresponding tumor marker concentration ranges were 2.0–51.5 ng/mL, 4.5–40.0 ng/mL, 2.0–29.5 ng/mL, and 8.5–36.5 ng/mL, respectively. The confidence detection range was wide and could meet the clinical application. When the concentration was too low (<0.9 ng/mL), the detection sensitivity could also decrease [43].

4. Conclusion

Based on the PCA method, the neighborhood rough set algorithm was introduced to improve it, and the tumor gene feature selection model (OPCA) was established. Furthermore, an immunobiosensor based on multicolor QDs was prepared and applied to the quantitative detection of potential gastrointestinal tumor markers. The results showed that the OPCA model can obtain fewer feature gene sets and improve the accuracy of sample classification. Intelligent immunobiosensors based on OPCA model and multicolor QDs had high specificity in the detection of gastrointestinal tumor markers. However, there were still some shortcomings in this study. In this study, only standard samples were tested, not clinical samples. In the future work, clinical samples of patients with gastrointestinal cancer will be further collected, and the intelligent immunobiosensor prepared in this study will be used to detect them, so as to

verify the clinical application value of the intelligent immunobiosensor based on multicolor QDs. In conclusion, the intelligent immunobiosensor based on machine learning algorithm and QDs has potential application value in gastrointestinal gene feature selection and tumor marker detection, which provides a new idea for clinical diagnosis of gastrointestinal tumors.

Data Availability

All data are fully available without restriction.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. Alanni, J. Hou, H. Azzawi, and Y. Xiang, "A novel gene selection algorithm for cancer classification using microarray datasets," *BMC Medical Genomics*, vol. 12, no. 1, p. 10, 2019.

[2] Y. Chen, Z. Zhang, J. Zheng, Y. Ma, and Y. Xue, "Gene selection for tumor classification using neighborhood rough sets and entropy measures," *Journal of Biomedical Informatics*, vol. 67, pp. 59–68, 2017.

[3] D. Yang and X. Zhu, "Gene correlation guided gene selection for microarray data classification," *BioMed Research International*, vol. 2021, pp. 1–11, 2021.

[4] M. Dashtban, M. Balafar, and P. Suravajhala, "Gene selection for tumor classification using a novel bio-inspired multi-objective approach," *Genomics*, vol. 110, no. 1, pp. 10–17, 2018.

[5] L. X. Zhang, H. Yan, Y. Liu, J. Xu, J. Song, and D. J. Yu, "Enhancing characteristic gene selection and tumor classification by the robust laplacian supervised discriminative sparse PCA," *Journal of Chemical Information and Modeling*, vol. 62, no. 7, pp. 1794–1807, 2022.

[6] G. Mantese, "Gastrointestinal stromal tumor," *Current Opinion in Gastroenterology*, vol. 35, no. 6, pp. 555–559, 2019 Nov.

[7] C. Karakas, P. Christensen, D. Baek, M. Jung, and J. Y. Ro, "Dedifferentiated gastrointestinal stromal tumor: recent advances," *Annals of Diagnostic Pathology*, vol. 39, pp. 118–124, 2019 Apr.

[8] A. Khoshnood, "Gastrointestinal stromal tumor - a review of clinical studies," *Journal of Oncology Pharmacy Practice*, vol. 25, no. 6, pp. 1473–1485, 2019 Sep.

[9] Y. Yin and B. Zhang, "[Clinical diagnosis and treatment of gastrointestinal stromal tumor: matching technological breakthrough with patient care]," *Zhonghua Wei Chang Wai Ke Za Zhi*, vol. 23, no. 9, pp. 852–857, 2020 Sep 25.

[10] A. Acharya, S. R. Markar, M. Matar, M. Ni, and G. B. Hanna, "Use of tumor markers in gastrointestinal cancers: surgeon perceptions and cost-benefit trade-off analysis," *Annals of Surgical Oncology*, vol. 24, no. 5, pp. 1165–1173, 2017 May.

[11] P. Das and V. Baloda, "Use of smooth muscle markers is better than the endothelial cell markers for identification of tumor venous invasion and extramural tumor deposits in gastrointestinal tract tumors," *Indian Journal of Pathology & Microbiology*, vol. 63, no. 1, pp. 3-4, 2020 Jan-Mar.

[12] D. P. M. Campos, J. G. Dórea, A. S. Galdino, Z. G. M. Lacava, and M. de Fatima Menezes Almeida Santos, "Carcinoembryonic antigen (CEA) and hepatic metastasis in colorectal cancer: update on biomarker for clinical and biotechnological approaches," *Recent Patents on Biotechnology*, vol. 12, no. 4, pp. 269–279, 2018.

[13] L. Deng, S. Guo, H. Li, X. You, Y. Song, and H. Su, "CA125, CEA, CA19-9, and heteroploid cells in ascites fluid may help diagnose peritoneal carcinomatosis in patients with gastrointestinal and ovarian malignancies," *Cancer Management and Research*, vol. 12, pp. 10479–10489, 2020 Oct 22.

[14] Y. Gao, J. Wang, Y. Zhou, S. Sheng, S. Y. Qian, and X. Huo, "Evaluation of serum CEA, CA19-9, CA72-4, CA125 and ferritin as diagnostic markers and factors of clinical parameters for colorectal cancer," *Scientific Reports*, vol. 8, no. 1, p. 2732, 2018 Feb 9.

[15] X. Li, B. Pasche, W. Zhang, and K. Chen, "Association of MUC16 mutation with tumor mutation load and outcomes in patients with gastric cancer," *JAMA Oncology*, vol. 4, no. 12, pp. 1691–1698, 2018 Dec 1.

[16] H. Yu, "Reference intervals for gastrointestinal tumor markers (AFP, CEA, CA199 and CA724) in healthy adults of Han nationality in Chongqing by Roche ECLIA system," *Scandinavian Journal of Clinical and Laboratory Investigation*, vol. 79, no. 7, pp. 484–490, 2019 Nov.

[17] F. Feng, Y. Tian, G. Xu et al., "Diagnostic and prognostic value of CEA, CA19-9, AFP and CA125 for early gastric cancer," *BMC Cancer*, vol. 17, no. 1, p. 737, 2017 Nov 9.

[18] B. Ghaderi, H. Moghbel, N. Daneshkhah, A. Babahajian, and F. Sheikhesmaeili, "Clinical evaluation of serum tumor markers in the diagnosis of gastric adenocarcinoma staging and grading," *Journal of Gastrointestinal Cancer*, vol. 50, no. 3, pp. 525–529, 2019 Sep.

[19] P. R. C. Dolscheid, S. Manekeller, B. G. Walgenbach et al., "Clinical performance of CEA, CA19-9, CA15-3, CA125 and AFP in gastrointestinal cancer using LOCI-based assays," *Anticancer Research*, vol. 37, no. 1, pp. 353–360, 2017 Jan.

[20] J. Wang, W. Liu, K. Parikh, and A. B. Post, "Alpha-fetoprotein-producing esophageal adenocarcinoma: a mimicker of hepatocellular carcinoma," *Clinical Journal of Gastroenterology*, vol. 10, no. 1, pp. 7–12, 2017 Feb.

[21] T. Nicol, C. Lefeuvre, O. Serri et al., "Assessment of SARS-CoV-2 serological tests for the diagnosis of COVID-19 through the evaluation of three immunoassays: two automated immunoassays (Euroimmun and Abbott) and one rapid lateral flow immunoassay (NG Biotech)," *Journal of Clinical Virology*, vol. 129, Article ID 104511, 2020 Aug.

[22] A. Moyano, E. Serrano-Pertierra, M. Salvador, J. C. Martínez-García, M. Rivas, and M. C. Blanco-López, "Magnetic lateral flow immunoassays," *Diagnostics*, vol. 10, no. 5, p. 288, 2020 May 8.

[23] B. Khlebtsov and N. Khlebtsov, "Surface-enhanced Raman scattering-based lateral-flow immunoassay," *Nanomaterials*, vol. 10, no. 11, p. 2228, 2020 Nov 10.

[24] H. Song, Y. Su, L. Zhang, and Y. Lv, "Quantum dots-based chemiluminescence probes: an overview," *Luminescence*, vol. 34, no. 6, pp. 530–543, 2019 Sep.

[25] S. Pandey and D. Bodas, "High-quality quantum dots for multiplexed bioimaging: a critical review," *Advances in Colloid and Interface Science*, vol. 278, Article ID 102137, 2020 Apr.

[26] Z. Lv, Y. Wang, J. Chen, J. Wang, Y. Zhou, and S. T. Han, "Semiconductor quantum dots for memories and neuromorphic computing systems," *Chemical Reviews*, vol. 120, no. 9, pp. 3941–4006, 2020 May 13.

[27] Y. Guo and J. Li, "MoS2 quantum dots: synthesis, properties and biological applications," *Materials Science and Engineering: C*, vol. 109, Article ID 110511, 2020 Apr.

[28] C. M. Feng, Y. Xu, J. X. Liu, Y. L. Gao, and C. H. Zheng, "Supervised discriminative sparse PCA for com-characteristic gene selection and tumor classification on multiview biological data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 10, pp. 2926–2937, 2019.

[29] S. Sayed, M. Nassef, A. Badr, and I. Farag, "A nested genetic algorithm for feature selection in high-dimensional cancer microarray datasets," *Expert Systems with Applications*, vol. 121, pp. 233–243, 2019.

[30] C. M. Feng, Y. Xu, J. X. Liu, Y. L. Gao, and C. H. Zheng, "Supervised discriminative sparse PCA for com-characteristic gene selection and tumor classification on multiview biological data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 10, pp. 2926–2937, 2019.

[31] H. Shariatifar, M. S. Hakhamaneshi, M. Abolhasani et al., "Immunofluorescent labeling of CD20 tumor marker with quantum dots for rapid and quantitative detection of diffuse large B-cell non-Hodgkin's lymphoma," *Journal of Cellular Biochemistry*, vol. 120, no. 3, pp. 4564–4572, 2019 Mar.

[32] C. Wang, F. Hou, and Y. Ma, "Simultaneous quantitative detection of multiple tumor markers with a rapid and sensitive multicolor quantum dots based immunochromatographic test strip," *Biosensors and Bioelectronics*, vol. 68, pp. 156–162, 2015 Jun 15.

[33] J. Kwon, S. W. Jun, S. I. Choi et al., "FeSe quantum dots for in vivo multiphoton biomedical imaging," *Science Advances*, vol. 5, no. 12, Article ID eaay0044, 2019 Dec 6.

[34] Y. Zhang, W. Ye, C. Yang, and Z. Xu, "Simultaneous quantitative detection of multiple tumor markers in microfluidic nanoliter-volume droplets," *Talanta*, vol. 205, Article ID 120096, 2019 Dec 1.

[35] G. Wang, Z. Li, and N. Ma, "Next-generation DNA-functionalized quantum dots as biological sensors," *ACS Chemical Biology*, vol. 13, no. 7, pp. 1705–1713, 2018 Jul 20.

[36] A. K. Jigyasu, S. Siddiqui, A. Jafri, M. Arshad, M. Lohani, and I. A. Khan, "Biological synthesis of CdTe quantum dots and their anti-proliferative assessment against prostate cancer cell line," *Journal of Nanoscience and Nanotechnology*, vol. 20, no. 6, pp. 3398–3403, 2020 Jun 1.

[37] H. Ahmadpour and S. M. M. Hosseini, "A solid-phase luminescence sensor based on molecularly imprinted polymer-CdSeS/ZnS quantum dots for selective extraction and detection of sulfasalazine in biological samples," *Talanta*, vol. 194, pp. 534–541, 2019 Mar 1.

[38] X. Li, Y. Wang, Q. Fu et al., "Plasmon-emitter hybrid nanostructures of gold nanorod-quantum dots with regulated energy transfer as a universal nano-sensor for one-step biomarker detection," *Nanomaterials*, vol. 10, no. 3, p. 444, 2020 Mar 1.

[39] Y. Wang, M. Yang, Y. Ren, and J. Fan, "Cu-Mn codoped ZnS quantum dots-based ratiometric fluorescent sensor for folic acid," *Analytica Chimica Acta*, vol. 1040, pp. 136–142, 2018 Dec 21.

[40] C. X. Liu, J. Zhao, R. R. Zhang et al., "Development and application of fluorescence sensor and test strip based on molecularly imprinted quantum dots for the selective and sensitive detection of propanil in fish and seawater samples," *Journal of Hazardous Materials*, vol. 389, Article ID 121884, 2020 May 5.

[41] B. D. Mansuriya and Z. Altintas, "Applications of graphene quantum dots in biomedical sensors," *Sensors*, vol. 20, no. 4, p. 1072, 2020 Feb 16.

[42] Z. Chen, P. Li, Z. Zhang et al., "Ultrasensitive sensor using quantum dots-doped polystyrene nanospheres for clinical diagnostics of low-volume serum samples," *Analytical Chemistry*, vol. 91, no. 9, pp. 5777–5785, 2019 May 7.

[43] Y. Upadhyay, S. Bothra, R. Kumar, A. Kumar Sk, and S. K. Sahoo, "Mimicking biological process to detect alkaline phosphatase activity using the vitamin B6 cofactor conjugated bovine serum albumin capped CdS quantum dots," *Colloids and Surfaces B: Biointerfaces*, vol. 185, Article ID 110624, 2020 Jan 1.

Research Article

Monitoring Technology of Abnormal Displacement of BeiDou Power Line Based on Artificial Neural Network

Jingbo Yang [iD], Yihan Chen, Jiarong Yu, Zheng Zhou, Yanna Guo, and Xingye Liu

Wuxi Guangying Group Co., Ltd., Wuxi 214000, Jiangsu, China

Correspondence should be addressed to Jingbo Yang; 41826059@xs.ustb.edu.cn

In the practice of power line engineering, navigation and positioning technology is often used in the fields of information collection and analysis, optimized line design, and deformation monitoring. Compared with traditional measurement technology, it has the characteristics of high precision and high reliability. In order to realize the measurement of abnormal displacement of power lines, improve the efficiency and quality of monitoring, and reduce the occurrence of faults, firstly, this study introduces the basic theory of artificial neural network (ANN). The core algorithm of the ANN-BP (back propagation) neural network has been improved. The improved algorithm is used to improve the BeiDou Navigation Satellite System (BDS). The improved and the unimproved BDS are used to solve the collected related data. The results show that the geometric dilution of precision (GDOP) values obtained by conventional BDS are small, all within the range of less than 4. After the introduction of the BP neural network into the system, the geometric space distribution of positioning satellites is improved, the GDOP is reduced, the reliability and availability of satellite positioning are enhanced, and the accuracy requirements are met. The accuracy of the measured data positioning results of the two systems has reached the cm level. There is not much difference between the processing results of the two modes. Among them, the Z direction accuracy has the largest difference, which is 2.5 cm. The introduction of the BP neural network has improved the spatial combination structure, and the positioning results in the three directions of X, Y, and Z are all better. From the perspective of root mean square (RMS), the RMS fluctuation of the simulation results obtained by observing the conventional BDS is large. The RMS value of BDS displacement based on the BP neural network is smaller, and the change is gentle. With the increase in the number of epochs and the increase in the number of simulations, its value is also more convergent. These data show that the quality of BDS observations based on the BP neural network is significantly better. These contents will effectively improve the monitoring accuracy and operational reliability of the system and have important practical significance and application value.

1. Introduction

The power line is one of the main facilities of the power system. Its normal and stable operation is the basis for ensuring the safety and trouble-free of the power system. Power towers and power transmission and transformation lines distributed in the wild may experience abnormal displacement and other phenomena due to the interference of external factors, resulting in serious accidents. Such accidents will have very adverse effects on the social economy and people's lives. The related research on abnormal displacement monitoring method of power lines has attracted extensive attention from scholars [1]. The purpose is to analyze the principle of the BeiDou dual-mode positioning system in the application of abnormal displacement monitoring of power lines, study the conversion technology of different reference systems, and unify the two-time systems and coordinate systems. The main sources of error are analyzed. Electromagnetic wave delay models for the ionosphere and troposphere are established. The satellite and receiver clock error calculation model eliminates errors through mathematical models such as double difference and improves the accuracy of observation data.

At present, there are two main early warning methods for transmission line status: the first one is the early warning method based on big data, deep learning, and other

computer technologies. Hubei Electric Power Co., Ltd. adopted the icing prediction model based on extreme learning machine and the icing prediction model based on support vector machine (SVM) to predict the icing of transmission lines. The experimental comparison results show that the prediction model based on SVM is better [2, 3]. The second is to divide the status level according to the evaluation result of the transmission line and take corresponding measures to carry out early warning according to the status level. The improved fuzzy comprehensive evaluation (FCE) is used for early warning of conductor galloping in transmission lines. Based on the optimal theory to optimize the weight, a comprehensive evaluation model is constructed in combination with the principle of fuzzy mathematics to improve the influence of the subjective weight of experts in the traditional fuzzy hierarchical evaluation. The rapid development of smart grids brings higher maintenance costs and greater scalability of transmission lines. An effective and safe power line monitoring system has become a bottleneck restricting the intelligentization of the power grid. In response to this problem, Fan et al. [4] proposed a new method for intelligent monitoring of power grids based on inspection robots, wireless sensor networks, and fault-tolerant sensor networks to achieve low cost, energy-saving, elastic, and remote monitoring. Through robotic fault-tolerant sensor networks, smart grids can detect faults in transmission lines and assess the operating state of the grid. Natural disasters and physical failures of overhead transmission lines have a severe impact on the grid, such as mechanical failures, power losses, line capacity reductions, and voltage drops. Jeyanthi et al. [5] pointed out that these adverse effects can be reduced by implementing an appropriate monitoring system. These methods have achieved certain results in the application process, but there are still some shortcomings and deficiencies. How to improve the accuracy of abnormal displacement monitoring of BeiDou power lines is still an issue that has been explored and discussed in academic circles.

In the practice of power line engineering, navigation and positioning technology is often used in the fields of information collection and analysis, optimized line design, and deformation monitoring. Artificial neural network (ANN) is introduced into BDS. Firstly, the basic theory of ANN is introduced. The core algorithm of ANN, the back propagation (BP) neural network, is improved. An improved algorithm is used to improve the BDS. The improved and unimproved BDS are used to solve the collected correlation data. The innovation lies in the application of ANN to the abnormal displacement monitoring technology of power BeiDou lines. The improved and unimproved BDS are used to solve the collected related data and achieve better monitoring accuracy. These conclusions provide a reference for the monitoring of abnormal displacement of power lines to realize the measurement of abnormal displacement of power lines, improve the efficiency and quality of monitoring, and reduce the occurrence of faults.

2. Research-Related Theories and Experimental Design

2.1. Line Abnormal Monitoring. Identification of abnormal data: the identification range of abnormal data of line status includes data information related to the line connected to the big data analysis system. The identified data are identified as the corresponding data state estimate. Line data identification should be carried out in two stages, namely, the data preprocessing stage and the model prediction stage. The data preprocessing stage is processed with the help of rules to eliminate abnormal situations caused by problems such as line connection, data association, and line correspondence. Such cases are exceptions caused by artificial reasons, and the abnormal state is constant before modification. The model prediction stage mainly uses machine learning algorithms to predict whether the line will be abnormal. The object of its prediction is the data itself, which has nothing to do with the overall grid [6–8]. The specific identification of the two stages is shown in Figure 1.

In Figure 1, in the preprocessing stage, the set preprocessing module is used to process the system's data to realize the function of eliminating invalid data. The content of abnormal data identification includes univariate data identification and cross-data identification. Univariate data identification is processed, including incomplete data, invalid data, and abnormal data quality. Cross-data identification needs to consider the overall network of the circuit, including the inability of line points to correspond to lines, the judgment of whether the connection point is floating, the position of the circuit breaker or knife switch does not correspond, and the equipment connection point is incomplete. The relationship is missing, or the filling is incorrect, etc. The model prediction stage, based on the machine learning model, predicts the state of the line for a period, including abnormal or normal. Model prediction must have a time advance, and a reasonable early warning time needs to be set to facilitate the staff's response, processing, and repair work. In the process of data mining, judging whether the data are abnormal can be processed from the data distribution. For example, the clustering algorithm can be used to judge the samples far from the cluster center as abnormal data. In a project, anomaly states are different from purely data-based anomalies in data mining. The abnormal early warning of the project is to facilitate the work of monitoring and maintenance personnel, to deal with abnormal lines in a timely manner, and to prevent the waste of manpower and physics caused by abnormal conditions. Electricity anomaly is used as the criterion for abnormality judgment. The abnormal situation identification can be processed by using the three-phase voltage value and the three-phase current value of electric power, and different processing can be performed according to the voltage level of the line. According to China's national network power data standard specification, there are some criteria for judging whether the power equipment is abnormal: for power system equipment of 35 kV and above, the absolute value of the positive and negative deviation between the measured data
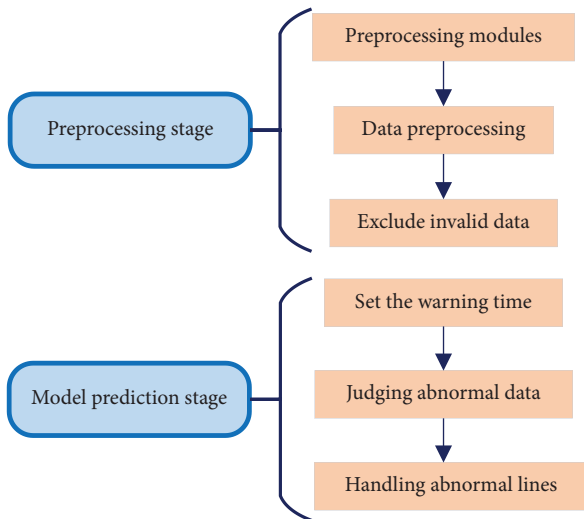
Figure 1: Two stages of line data identification.

Figure 2: Setting of early warning preliminary conditions.

value and the rated value shall not exceed 10%. The deviation between the measured value and the rated value of the power system equipment data of 10 kV and below shall not exceed 7%. Since the voltage levels of different lines and the same line at different times will be changed and adjusted, the data of different voltage levels should be processed separately during the processing.

Real-time early warning of abnormal state: a high-performance monitoring system should have the function of early warning in minutes. The alarm frequency of the system is consistent with the data collection frequency, which is performed once every minute. The prediction period is set to 30 minutes, that is, to predict whether the line will be abnormal after 30 minutes. The reason for being set to 30 minutes comprehensively considers the balance between the prediction accuracy and the response time and processing time of the staff after the warning [9]. If the prediction time is too short, although the accuracy rate can be improved, the processing time for the staff is short, and the staff may not be able to take timely measures to troubleshoot the fault. On the other hand, if the setting time is too long, although there is enough time for processing, the prediction accuracy is low, and the false-positive rate and coverage rate will be affected. In addition, it should be possible to set early warning rules, and for the prediction set with an abnormal prediction probability greater than the threshold, only the set abnormal situation is displayed. Based on generalized early warning, the customization of early warning objects is allowed; that is, monitoring personnel are allowed to set early warning conditions [10]. The preliminary conditions are set as shown in Figure 2.

In Figure 2, the first step in setting the preliminary warning conditions is to allow the blacklist to be set. The actual use object of the prediction system is the monitoring staff of various places. They have a better understanding of the warning line and are more aware of the line maintenance and stoppage time. So, they are allowed to set a blacklist. After the prediction results are filtered out of the lines in the blacklist, other abnormal lines are aggregated and displayed
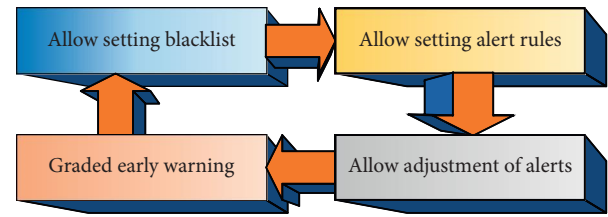
to the user. The second step is to allow the setting of alert rules. Unlike a blacklist that is specific to each line, an alert rule is a collection of several alert conditions. If the warning condition is more than four times within five minutes, the warning level is set to A-level warning. If the number of warnings is greater than two and less than 4, it is set as a B-level warning. The first warning is set as a C-level warning. Such warning rules can classify warnings and set higher levels for more serious warnings, which is convenient for users to deal with in a timely manner. The third step is to allow adjustment of the early warning value. The prediction result of the prediction model is the corresponding abnormal probability. For the line whose abnormal probability is greater than a certain threshold, an early warning is performed, and the initial value of the early warning value can be set to the optimal value. Additionally, dynamic adjustment of the threshold value is allowed. There is a trade-off between forecast accuracy and forecast coverage. The fourth step is graded early warning. Like the second step, early warning is carried out according to the conditions and levels of the predicted probability. Like the second step, the early warning is divided into conditions and levels according to the predicted probability. The difference is that the second step allows user-defined rules to be limited. The graded early warning refers to the predicted abnormal probability value of the line, and the line with a higher abnormal probability value should be dealt with first.

Statistical analysis of historical abnormal data: the abnormal prediction of line status is classified and archived according to region and time, which is convenient for data analysts to conduct statistical analysis. The statistical analysis of historical data is mainly to compare and analyze the predicted state of the transmission line in the statistical period between the measured state value and the corresponding state predicted value [11]. The forecast situation and the monitoring situation are analyzed. Statistical analysis results will be displayed in the form of lists, graphs, curves, etc. The results of classifying the forecast data by region, equipment type, time period, etc., are shown in Figure 3.

Figure 3 shows (1) a summary analysis by region. In order to facilitate the comparison between regions, the forecast should be classified and archived according to the location of the line, the rational use of the line between different regions should be analyzed, and the transmission pattern should be planned and adjusted. Whether the use of transmission lines is reasonable and whether the losses caused by abnormality are minimized are closely related to the cost of transmission lines. A more reasonable situation is
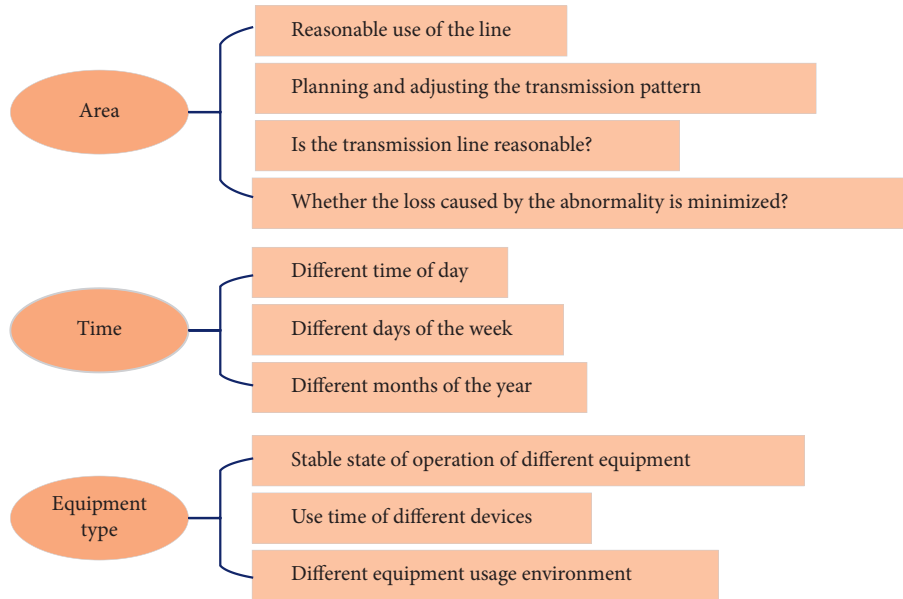
FIGURE 3: Subtotal summary of forecast data.

to coordinate the transmission cost between each line and the cost of waste caused by an abnormality. In order to facilitate optimal planning and processing, abnormal prediction results and actual measurement results should be classified and archived according to regions to facilitate future use. (2) Summarize and analyze according to time. Line anomalies may vary with different time periods of the day, different days of each week and month, different months and seasons of each year, etc., which requires the forecast results to be aggregated by time. A high probability of anomalies usually accompanies high-power transmission. Transmission power is usually strongly correlated with time period, so it is necessary to summarize by time. Combined with location summary, dispatch to high-consumption areas when the line is idle, reducing the cost loss caused by line anomalies caused by high-power transmission. (3) Summarize and analyze according to the equipment type. Different types of equipment are related to the stable operation state, usage time, usage environment, and abnormal occurrence. Summarize according to equipment type to facilitate statistical analysis of abnormal conditions caused by a different equipment, and use different types of equipment under reasonable conditions.

Analysis of the cause of abnormal line status: the purpose is to provide advanced analysis, diagnosis, and processing services for abnormal data, so that relevant staff have time to do corresponding processing in advance and reduce the loss of manpower and physical costs caused by abnormal line status. In order to explain the abnormal situation, the abnormal situation needs to be explained, and the reason for the abnormal situation of the predicted line needs to be listed. The reason for this is to facilitate troubleshooting by staff. From the perspective of model prediction, not only the prediction accuracy of the model is required to be high, but also the selection of an algorithm with strong interpretability [12]. For some models, such as

deep learning, the research results in data prediction are better, and they do not need to make explicit explanations and have better applicability.

2.2. Overview of ANN. Simplified neuron mathematical model: the artificial neuron network is an information processing system composed of the structure and function of the physiological real human brain neural network, as well as some theoretical abstraction, simplification, and simulation of some basic characteristics. It is an adaptive nonlinear dynamic system composed of many neurons through extremely rich and perfect connections.

The brain's neural network is a network composed of many highly interconnected neurons. It realizes the processing and storage of information through the interaction between neurons in the network. The brain neural network imitates the structure and function of the brain neural network and connects artificial neurons into a network according to certain structures and rules so that the connection weight of each neuron in the network changes according to certain rules to realize the learning or recognition of input patterns.

The brain neural network is a network composed of many highly interconnected neurons with statistical regularity. Due to the difficulty of physical implementation and the simplicity of calculation, the number of neurons that make up ANN is far less than that of a brain neural network. ANN is formed completely according to certain rules. Each neuron in ANN has the same structure, and in general, the actions of all neurons are synchronized in time and space.

The topological structure of ANN is generally divided into two types according to the flow direction of information: feedforward type and feedback type. ANN is also divided into a forward network and a feedback network. Some ANN has the same topology but has different functions and properties. This is because they have different study rules and

work rules. The learning and working rules are the dynamic evolution rules of the connection weights between neurons in the network. In short, two main factors determine the properties of an ANN: one is the topology of the network; the other is the learning and working rules of the network. The formalized structural model of the artificial neuron is shown in Figure 4 [13].

In Figure 4, $x_1, x_2, x_3, \ldots, x_n$ are input signals. Assumption: there are n neurons interconnected here, and the first neuron is used as the object. It inputs information to all other neurons $j (j = 1, 2, \ldots, n)$. The connection weight from the jth neuron to the ith neuron is denoted as ω_{ij}.

Taking the quasi-linear unit model in the neuron model as an example, it uses continuous information as input and output. The biggest feature of this model is that the output function $f(x)$ is in the form of

$$f(x) = \frac{1}{1 + \exp(-x + \theta_i)}. \tag{1}$$

The total input μ_i of this model neuron i is given by

$$\mu_i = \sum_{j=1}^{n} \omega_{ij} x_j - \theta_i. \tag{2}$$

In (2), μ_i is substituted into (3) as a variable x, and the output of neuron i can be calculated, as shown in

$$y_i = f_i(\mu_i). \tag{3}$$

The output value y_i is also a continuous value. Quasi-linear unit models are widely used in BP networks [14].

2.3. BP Neural Network. Multi-layer feedforward network, also known as multi-layer feedforward network or multi-layer perceptron network, is one of the typical ANN models. It is also the earliest, most researched, and most widely used type of ANN model in pattern recognition and classification. The multi-layer forward network in which the back propagation algorithm is used for training is called a back propagation neural network, or BP network for short [15, 16]. In practical application, the BP network reflects the essential part of ANN at this stage. The main application scope of the BP network is shown in Figure 5.

The training process of the BP algorithm is shown in Figure 6.

Forward propagation of the working signal: the input signal is transmitted from the input layer to the output layer through the hidden unit, and a signal is generated at the output end. This is the forward propagation of the working signal. The weights of the network are fixed during the forward transmission of the signal. The state of neurons in each layer only affects the state of neurons in the next layer. If the desired output cannot be obtained at the output layer, the error signal is transferred to back propagation.

Error signal back propagation: the difference between the actual output and the expected output of the network is the error signal. The error signal starts from the output and propagates forward layer by layer. This is the back propagation of the error. During the back propagation of the error
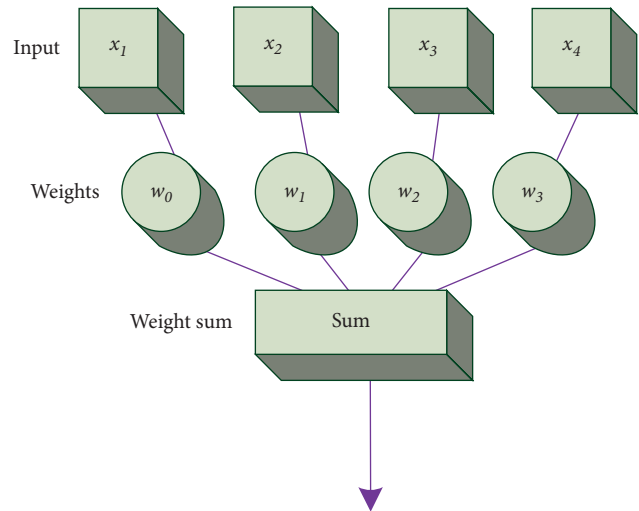
Figure 4: Mathematical model of artificial neuron.

signal, the network weights are adjusted by the error feedback. The actual output of the network is closer to the expected output through the constant correction of the weights [17].

The back propagation process of the error and the forward propagation process of the signal together constitute the learning process of the BP algorithm. During the forward propagation of the signal, each neuron of the input layer is provided with input samples. The net inputs and outputs of the output layer and each hidden layer are computed. Further, the prediction result of the neural network is calculated. If there is a large error between the calculated prediction result and the expected output, the back propagation of the error is started. During the back propagation of the error, the input error is reversed in the output layer. The errors are back-propagated, and the weights are updated based on the errors at each layer. During this process, the weights will be continuously updated and adjusted. The learning process ends when the learning process reaches a preset number of cycles or the neural network output error is less than the specified threshold.

A typical BP network is a three-layer feedforward hierarchical network, that is, an input layer, a hidden layer (also called an intermediate layer), and an output layer. A full connection is implemented between each layer, as shown in Figure 7.

In Figure 7, structurally speaking, the BP network is divided into an input layer, a hidden layer, and an output layer. Nodes at the same level are not related, and neurons at different levels propagate from front to back. Among them, the BP network corresponding to several nodes in the input layer can perceive several inputs. The output layer contains several nodes, and the BP network will have several output data. The number of nodes in the hidden layer needs to be adjusted or set according to the actual situation. The more the nodes in the hidden layer, the higher the accuracy of the result and the more time required [18]. Let the input layer be M. That is, there are M input signals, any of which is represented by m. The hidden layer is J. That is, there are J
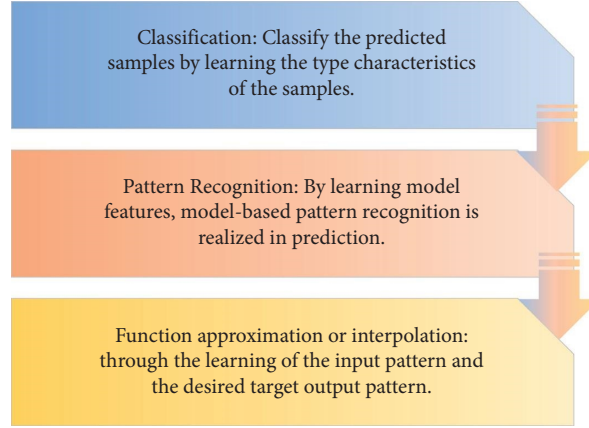
FIGURE 5: The main application scope of BP network.
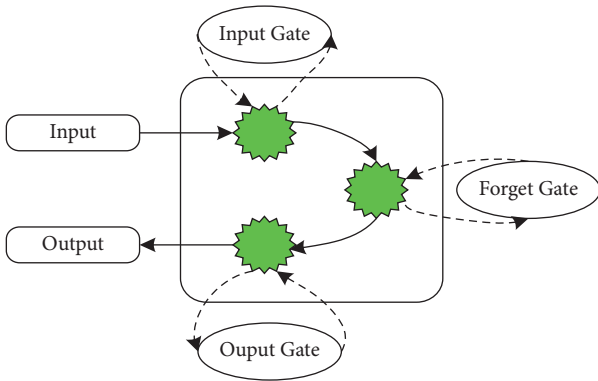
FIGURE 6: The training process of the BP algorithm.
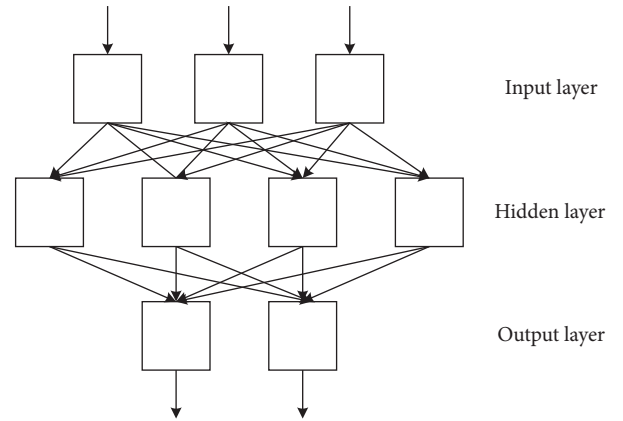
FIGURE 7: Schematic diagram of BP neural network.

neurons, and any neuron is represented by j. The output layer is P. That is, there are P output neurons, and any neuron is represented by p. The synaptic weights of the input layer and the hidden layer are denoted by W_{mj}. The synaptic weights of the hidden layer and the output layer are denoted by W_{jp}. The input of the neuron is denoted by u, and the excitation output is denoted by v. The superscript of u and v represents the layer, and the subscript represents a neuron in the layer. It is assumed that the excitation function of all neurons uses the sigmoid function. The specific function formula between each layer is as follows:

The output function of the node in the hidden layer is shown in

$$b_r = f\left(\sum_{r=1}^{m} W_{ir} \cdot a_i + T_r\right)(r = 1, 2, \ldots, u). \qquad (4)$$

The output function of the node in the output layer is shown in

$$c_j = f\left(\sum_{r=1}^{u} V_{ir} \cdot b_r + \theta_j\right)(j = 1, 2, \ldots, n). \qquad (5)$$

Among them, the input layer node is a_i. The connection weight between the input layer node and the hidden layer node b_x is W_{ir}. The connection weight between the hidden

layer node b_x and the output layer node c_j is V_{ir}. T_r is the threshold of hidden layer nodes. θ_j is the threshold of the output layer node. $f(\cdot)$ is the transfer function, and usually the sigmoid transfer function is selected, as shown in

$$P = \left(1 + e^{-x}\right)^{-1}. \qquad (6)$$

The maximum value approaches 1, and the minimum value approaches 0. Typically, 0.5 is chosen as the threshold. Variables that are completely correlated are recorded as 1, and variables that are not correlated are recorded as 0. If the P value obtained by the operation is greater than 0.5, it indicates that the variable is highly correlated. If the obtained P value is less than 0.5, it indicates that the correlation between variables is low [19].

The specific BP network learning process is as follows:

Step 1 is to randomly assign a smaller value to W_{ir}, T_x, V_{ir}, and θ_j.

Step 2 is to perform the following operations for each pattern $(A^{(k)}, A^{(k)})$ $(k = 1, 2, \ldots, p)$:

(1) The value of $A^{(k)}$ is input to the input layer node; that is, it becomes the input layer node activation value a_i and is calculated forward in turn.

(2) The error calculation between the output c_j of the output layer node and the expected output value $c_j^{(k)}$ is shown in

$$d_j = c_j \cdot \left(1 - c_j\right) \cdot \left(c_j^{(k)} - c_j\right). \qquad (7)$$

(3) The hidden layer nodes have assigned the error inversely, as shown in

$$e_x = b_x \cdot \left(1 - b_x\right) \cdot \left(\sum_{j=1}^{n} V_{rj} \cdot d_j\right). \qquad (8)$$

(4) The connection weight W_{ir} between the input layer and the hidden layer nodes and the threshold T_r of the hidden layer nodes are adjusted:

$$\begin{aligned} W_{ir} &= W_{ir} + \beta \cdot a_i \cdot e_r, \\ T_r &= T_r + \beta \cdot e_r \, (0 < \beta < 1). \end{aligned} \qquad (9)$$

(5) Step 2 is repeated until the error E_{AV} becomes sufficiently small for $j = 1, 2, \cdots, n, k = 1, 2, \cdots, p$.

$$E_{AV} = \frac{1/2 \sum_{k=1}^{p} \sum_{j=1}^{n} \left(c_j^{(k)} - c_j\right)^2}{P}. \qquad (10)$$

Among them, E_{AV} is the training objective function. After repeated training, the error E_{AV} meets the accuracy required in the specific problem.

Two serious shortcomings of the BP network algorithm are the slow convergence speed and the existence of local minimum points in the objective function. These deficiencies affect the practical application of this network in many aspects. Therefore, many scholars have conducted extensive research on the learning algorithm of BP network and proposed some methods to improve the BP algorithm. The more common method is to add a momentum term. In the BP algorithm, the selection of the learning step η is very important. The larger the value of η, the faster the network converges, but too fast will cause oscillation. Although the small value of η can avoid oscillation, the convergence speed will be slow [20]. The easiest way to resolve this contradiction is to add a momentum term, as shown in

$$\Delta \omega_{mj}(n) = a \Delta \omega_{mj}(n-1) + \eta \delta_j(t) v_m(t). \qquad (11)$$

Adding the momentum term to the BP algorithm can not only fine-tune the correction of the weights, but also prevent the learning from falling into a local minimum. When deriving the BP algorithm, the learning parameter η is assumed to be constant. In fact, the η corresponding to different ω_{mj} is different.

3. Design of the Abnormal Displacement Detection Experiment of the Line

The purpose is to implement a BDS based on BP neural network. It is used to monitor abnormal displacement problems in power lines. Geometric dilution of precision (GDOP) is used to measure the positioning error in the BeiDou positioning algorithm. The localized error is proportional to the size of the GDOP [21]. The BDS based on BP neural network and the conventional BDS are used to solve the satellite data. The results of the two solutions are compared. The main comparison items included are shown in Figure 8.

The results of the two positioning algorithms are analyzed by the difference between the positioning solution value and the actual value. The experiment performs static differential positioning on the data received by BDS based on BP neural network. In order to analyze the relative positioning accuracy of the algorithm more intuitively, the weighted least squares method is used for parameter estimation in the analysis process. Secondly, the GDOP of BDS is analyzed, taking the data collected by BeiDou in the Asia-Pacific region on April 1, 2019, as an example, a total of 160 epochs. The results are compared with the precise coordinates of the International GNSS Service (IGS) stations provided by the Scripps Orbit and Permanent Array Center (SOPAC) [22, 23].

The positioning accuracy of the two algorithms is compared. According to the algorithm flow of conventional BDS phase relative positioning and BDS phase differential positioning based on BP neural network, and using the data collected in the Asia-Pacific region, the epoch interval is set to 30 s, and the observation time is April 1, 2013., 28800 epochs. The positioning results of the two algorithm systems are obtained, and the positioning accuracy of the two positioning systems is analyzed and compared.

Statistical analysis of the accuracy of positioning results: root mean square (RMS) is the probability statistics of the deviation between the position of the monitoring point and its real position obtained by BDS modeling. Its value is not affected by the observation conditions. Analysis of the RMS can estimate the positioning accuracy of the positioning system [24]. RMS is defined as

$$\text{RMS} = \sqrt{\frac{V^T P V}{n - k}}. \qquad (12)$$

Among them, V^T represents the residual size of the observed value. k is the difference between the total number of observations and the number of unknowns. P represents the weight of the observation. When P is determined, the size of the error is related to the residuals of the observations. The unit weight error increases as the residual increases. The RMS can be used to analyze the goodness of the accuracy of the positioning results. The smaller the value, the better the quality of the observations and the higher the accuracy. The time period with better observation conditions at the Shanghai station is selected, and the simulation is performed 200 times with a sampling interval of 1 s. In both modes, the RMS of the positioning results of the monitoring station in the three directions of X, Y, and Z is recorded.

Neural network parameter design: the standard input of the neural network with the standard activation function is taken. The output data are limited to the range [0, 1]. The parameters (X, Y) in practical engineering applications
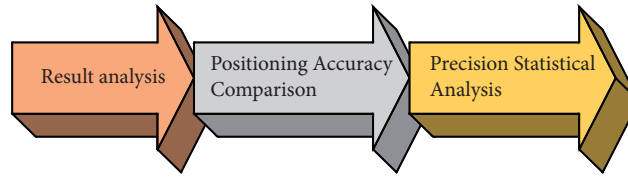
FIGURE 8: Comparison of the two algorithms.

have very large values and need to be converted into values in the [0, 1] interval. In addition, the region where the output is close to 0 or 1 is the saturation region of the network. Therefore, the output data range can be set to [0.2, 0.8] or [0.1, 0.9] to avoid the saturation region of the network.

Theoretically, GPS elevation fitting is a mapping from Rn ($n = 2$, representing the plane coordinates or geodetic coordinates of the GPS point) to Rm ($m = 1$, representing the abnormal or normal elevation of the GPS point). The classic BP neural network structure is a three-layer structure: input layer, hidden layer, and output layer. Therefore, the ideal BP neural network structure should be the structure of $A \times M \times B$. If the input layer takes two and the output layer takes 1, then the model $2 \times M \times 1$ is the best choice. Among them, in the network construction test, M is first selected as 10. Because ten known points are selected in this engineering test, ten is selected, the $2 \times 10 \times 1$ network is not redundant matching for known data, and ideal results can be obtained. For this purpose, a $2 \times 10 \times 1$ network structure is chosen.

4. Experimental Results

Analysis of positioning results: in the experiment, static differential positioning is performed using the data collected by the BDS/GPS dual-mode system receiver developed by ComNav Technology Ltd. In order to analyze the positioning accuracy of GPS/BDS dual-mode carrier phase static relative positioning more intuitively, the weighted least squares method is used to estimate the parameters in the analysis process. BDS/GPS dual-mode systems require dual-mode receivers to have good observation conditions for satellites. When the number of satellites is less than 5, the dual-mode positioning solution cannot be performed. When various factors that affect the observation appear, the number of visible satellites becomes smaller, which in turn reduces the positioning accuracy. Assumption: the altitude angle of the satellite is 15°. In order to minimize the influence of error sources such as receiver clock error, ionospheric delay, and tropospheric delay, short baseline data are discussed, with a baseline length of about 1 km.

Firstly, BDS and GPS satellites are analyzed. On April 1, 2018, from 12:00 to 13:20, the data collected by BDS and GPS in the Shanghai station in the Asia-Pacific region were taken as an example, with a total of 160 epochs. The number of visible satellites of BDS and GPS is similar. BDS is about 7–10, GPS is about 6–10, and GPS has more visible satellites than BDS. This is because the current BDS satellite launch volume has not reached the full planned number of GPS

satellites. The GDOP values of the two algorithms are shown in Figure 9.

Figures 9(a) and 9(b) show the simulated GDOP values with a different BDS positioning. The GDOP values obtained by conventional BDS are small, all in the range of less than 4. The geometric space distribution of positioning satellites is improved after the introduction of the BP neural network. The reduction of the GDOP value enhances the reliability and availability of satellite positioning and meets the accuracy requirements. The BDSGDOP value based on the BP neural network is smaller than that of the conventional system, and the positioning accuracy is better.

The solved data results are compared with the exact coordinates of the IGS station. The localization result of BDS short baseline vector based on BP neural network is shown in Figure 10.

In Figure 10, X represents the horizontal displacement data on the plane in the three-dimensional space, Y represents the vertical displacement data on the plane, and Z represents the vertical displacement data. The unit of solution value and reference value is m, and the unit of difference value is cm. The BDS displacement monitoring method based on BP neural network is basically consistent with the coordinate values obtained by using the IGS integrated track. The effect of the positioning results is more ideal with the continuous improvement of BDS and the further improvement of the precision of the corresponding precision track products.

Comparison of conventional BDS and improved BDS positioning accuracy: according to the algorithm flow of single GPS carrier phase relative positioning and BDS/GPS dual-mode carrier phase differential positioning, the positioning programs are written, respectively. Likewise, the data collected at the Shanghai station in the Asia-Pacific region are used. The epoch interval is set to 30 s. The observation time is April 1, 2018, with 2880 epochs throughout the day. In the navigation and positioning results, the accuracy statistics obtained by BDS based on BP neural network and conventional BDS are compared, as shown in Figure 11.

In Figure 11, ROU represents conventional BDS. The accuracy of the measured data positioning results of the two systems has reached the cm level. The difference is small between the processing results of the two modes. Among them, the Z direction accuracy has the largest difference, which is 2.5 cm. The introduction of the BP neural network has improved the spatial combination structure, and the positioning results in the three directions of X, Y, and Z are all better.

In order to better compare and analyze the positioning accuracy of the two systems, the error variance is obtained by the mean of the deviation of the positioning results, which is used as one of the indicators to reflect the accuracy of the

(a)

(b)
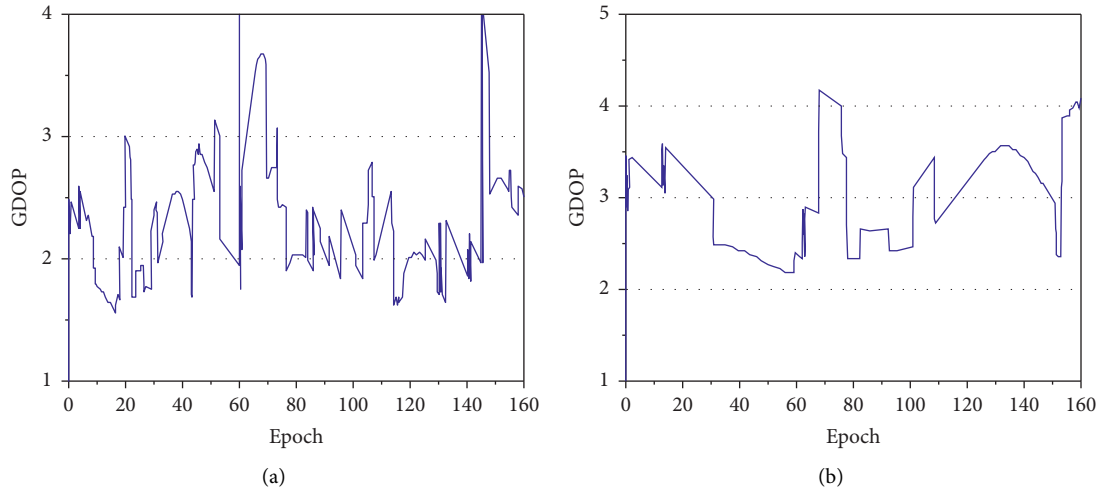
FIGURE 9: GDOP values of the two algorithms. (a) Conventional BDSGDOP value; (b) BDSGDOP value based on BP neural network.
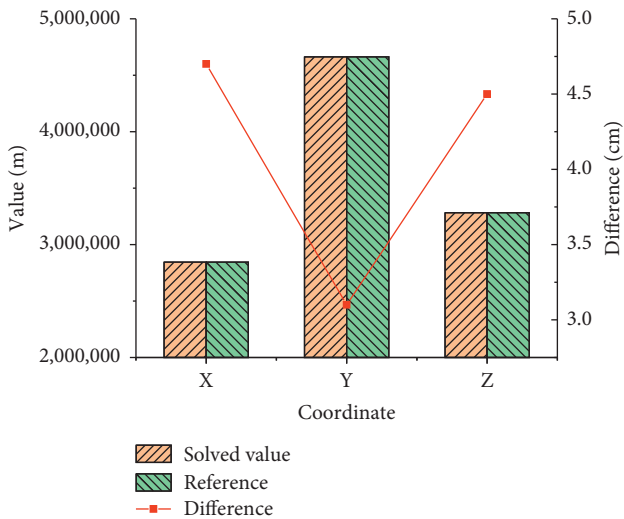
FIGURE 10: Comparison of positioning results and reference values.

FIGURE 11: Comparison of the positioning results and reference values of the two algorithms.

FIGURE 12: Analysis of error variance.

positioning results. The numerical statistics results are shown in Figure 12.

In Figure 12, compared with the conventional BDS, the variance value of the BDS based on the BP neural network is smaller, indicating that the phenomenon of poor satellite geometric distribution is less, the static differential positioning result obtained by the least squares method has less fluctuation, jump is small, and the results converge well. The positioning accuracy of the system is improved compared with the conventional system.

The RMS of the positioning results of the monitoring station in the X, Y, and Z directions in the two modes is shown in Figure 13.

Figures 13(a)–13(c) are the displacement errors of the coordinates of the monitoring station in the X, Y, and Z directions when the two systems are used for positioning, respectively. When the simulation parameters are the same and the operation is normal, different satellite systems are

(a)

(b)

(c)
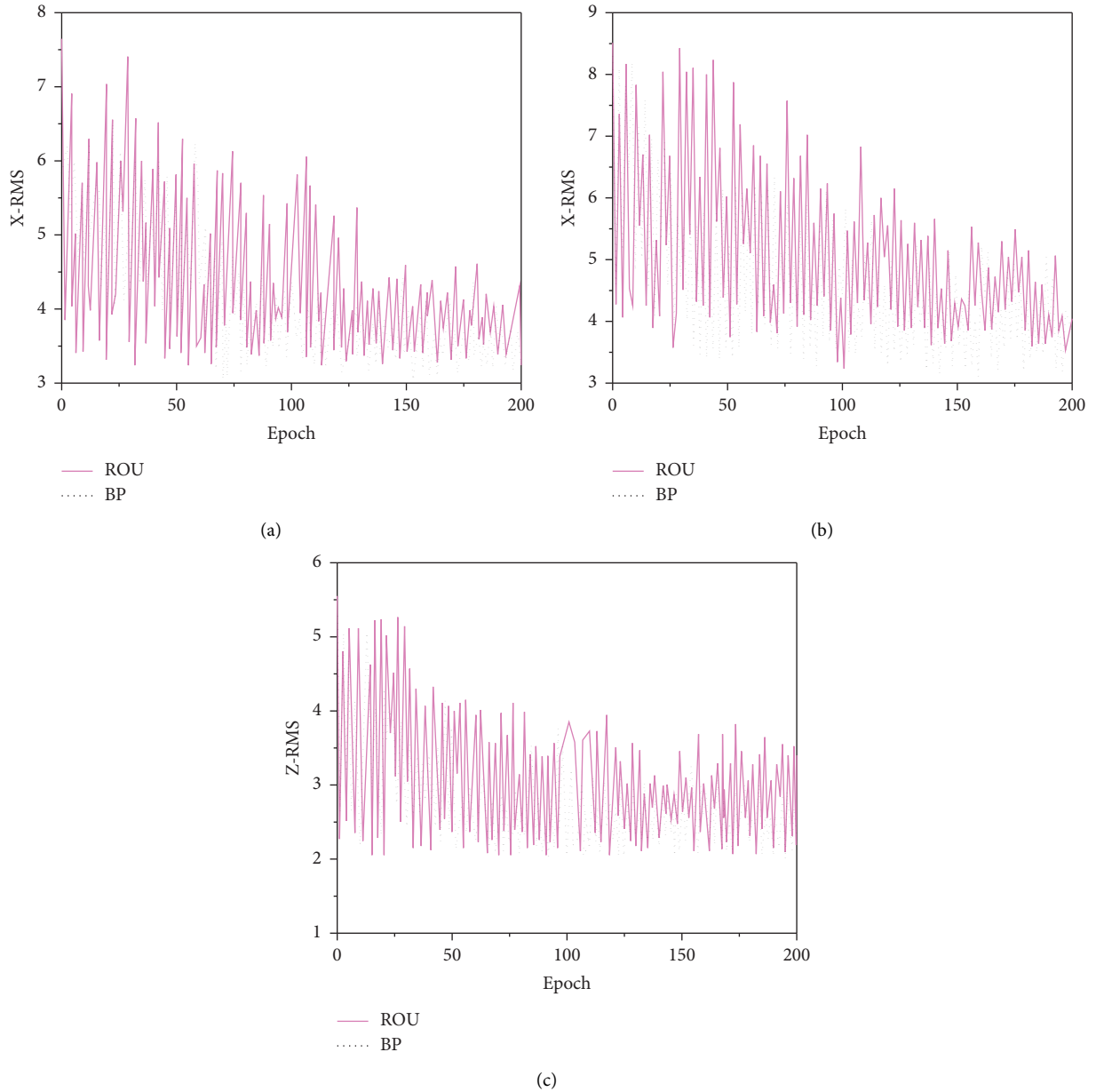
Figure 13: Positioning RMS of the two systems in different directions. (a) RMS in *X* direction; (b) RMS in *Y* direction; (c) RMS in *Z* direction.

tracked, and the RMS of the positioning result obtained by observing the conventional BDS fluctuates greatly. The RMS value of BDS displacement based on the BP neural network is smaller, and the change is gentle. As the number of epochs increases, the number of simulations increases, and its value becomes more convergent. Additionally, the RMS value obtained by the system operation is smaller than that of the conventional system, indicating that the BDS observation quality based on BP neural network is better.

The RMS means of the positioning results of the two systems is shown in Figure 14.

In Figure 14, the RMS average of the BDS localization results based on the BP neural network is smaller than that of the conventional BDS. Among them, the RMS of the positioning result error of the former is not more than 4 cm,

which reflects that the former has effectively improved the positioning data processing method of the latter. The positioning results obtained by this system are more stable and more reliable. From the point of view of mathematical statistics, the error fluctuation range of BDS positioning results based on BP neural network is relatively small, and it is better in positioning accuracy.

Bao et al. [25] used wavelet analysis theory and proposed a wavelet threshold filtering method based on empirical mode decomposition to detect overhead high-voltage transmission line faults. For overhead high-voltage transmission line faults, it should be noted that there are three types of noise: Gaussian white noise interference, periodic narrowband interference, and impulse interference. The appearance of these noise signals will increase the leakage
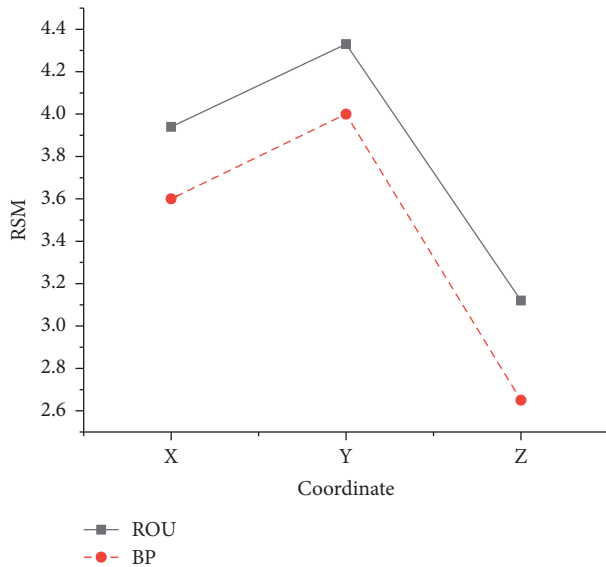
Figure 14: Statistics and comparison of RMS values.

current error of high-voltage transmission lines, thus interfering with fault detection. Therefore, the interference noise on-site must be eliminated or suppressed first, and the positioning result obtained by the system is more stable and more reliable. From the point of view of mathematical statistics, the error fluctuation range of BDS positioning results based on BP neural network is relatively small, and it is better in positioning accuracy.

5. Conclusion

Power lines are the most important link in the transmission of electrical energy from power grids to users. In order to reduce the failure of power lines and ensure the safety of people's lives and property, in addition to strengthening protection measures from personnel management, technically, it also puts forward higher requirements for abnormal displacement monitoring of power lines. It is urgent to develop a reliable and stable abnormal displacement monitoring and positioning system. The satellite positioning system performs positioning, timing, and monitoring of users on the Earth through the time, position, speed, and other signals broadcast by satellites.

BP neural network is used to optimize the BDS. The improved BDS and the unimproved conventional BDS are used to solve the collected correlation data. The accuracy of the solutions of the two systems is compared by designing experiments. The experimental results show that the GDOP values obtained by conventional BDS are small, all within the range of less than 4. After the introduction of the BP neural network, the geometric space distribution of the positioning satellites is improved, the GDOP is reduced, the reliability and availability of the satellite positioning are enhanced, and the accuracy requirements are met. The accuracy of the measured data positioning results of the two systems has reached the cm level. There is not much difference between the processing results of the two modes. Among them, the Z

direction accuracy has the largest difference, which is 2.5 cm. The introduction of BP neural network has improved the spatial combination structure, and the positioning results in the X, Y, and Z directions are all better. The RMS average of BDS localization results based on the BP neural network is smaller than that of conventional BDS, and the RMS of the former's localization result error does not exceed 4 cm. This reflects that the former has effectively improved the positioning data processing method of the latter. From the perspective of RMS, the simulation results obtained by conventional BDS have large fluctuations in RMS. The RMS value of BDS displacement based on the BP neural network is smaller, and the change is gentle. With the increase in simulation times, its value is more convergent, indicating that the quality of BDS observation based on the BP neural network is better. Due to limited energy, the study did not consider the monitoring accuracy of the system under an abnormal climate. In the future, this aspect will be further explored.

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.

References

[1] K. S. Umesh, "Performance Analysis of Power line Fault not touching the ground," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 8, pp. 1538–1543, 2019.

[2] L. Gang, Q. Chen-guang, C. Shuai, and W. Ya-O, "Fault warning method based on extreme learning regression and fuzzy reasoning," *E3S Web of Conferences*, vol. 236, no. 1, Article ID 04008, 2021.

[3] A. Yamanaka, N. Nagaoka, and Y. Baba, "Circuit model of an overhead transmission line considering the TEM-mode formation delay," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 16, no. 6, pp. 888–895, 2021.

[4] F. Fan, G. Wu, M. Wang, Q. Cao, and S. Yang, "Robot delay-tolerant sensor network for overhead transmission line monitoring," *Applied Sciences*, vol. 8, no. 6, 2018.

[5] C. Jeyanthi, H. H. Sait, K. Chandrasekaran, and C. C. Columbus, "A communication model framework for electric transmission line monitoring using artificial bee colony algorithm," *Journal of Circuits, Systems, and Computers*, vol. 29, no. 4, Article ID 2050063, 2019.

[6] B. Li, R. Zhao, J. Lu, W. Guo, and Y. Liu, "Control on abnormal data overflow of distribution network management platform," *Journal of Physics: Conference Series*, vol. 1748, no. 3, Article ID 032064, 2021.

[7] Z. Li, B. Sui, X. Chen, J. Zhao, and J. Tian, "Research on multi-node frame early warning system of power grid based on abnormal data extraction," *Journal of Physics: Conference Series*, vol. 1654, no. 1, Article ID 012020, 2020.

[8] H. Long, L. Sang, Z. Wu, and W. Gu, "Image-based abnormal data detection and cleaning algorithm via wind power curve,"

IEEE Transactions on Sustainable Energy, vol. 11, no. 2, pp. 938–946, 2020.

[9] R. Luo, Z. Tan, H. Gao, R. Liang, and S. Feng, "Research on spacecraft abnormal state detection method based on telemetry data," *Journal of Physics: Conference Series*, vol. 1213, no. 4, Article ID 042065, 2019.

[10] G. Al-Sinbol and M. G. Perhinschi, "Development of an artificial immune system for power plant abnormal condition detection, identification, and evaluation," *International Review of Automatic Control (IREACO)*, vol. 10, no. 3, pp. 218–228, 2017.

[11] Z. Pan, W. Ji, Y. Chen, L. Dai, and J. Zhang, "Anomaly data management and big data analytics: an application on disability datasets," *International Journal of Crowd Science*, vol. 2, no. 2, pp. 164–176, 2018.

[12] A. Mahyar, S. Alireza, J. Mahmood, and G. S. Seyyed, "Fault classification and fault area detection in GUPFC-compensated double-circuit transmission lines based on the analysis of active and reactive powers measured by PMUs," vol. 169, Measurement, Prepublish, Article ID 108499, 2021.

[13] G. L. Foresti, "Editorial: from pioneering artificial neural networks to deep learning and beyond," *International Journal of Neural Systems*, vol. 31, no. 5, Article ID 2103004, 2021.

[14] S. Cruz, A. Paulino, J. Duraes, and M. Mendes, "Real-time quality control of heat sealed bottles using thermal images and artificial neural network," *Journal of Imaging*, vol. 7, no. 2, p. 24, 2021.

[15] Y. Deng, L. Qiao, J. Zhu, and B. Yang, "Mechanical performance and microstructure prediction of hypereutectoid rail steels based on BP neural networks," *IEEE Access*, vol. 8, no. 1, Article ID 41905, 2020.

[16] C. Senthilkumar and R. K. Gnanamurthy, "A Fuzzy clustering based MRI brain image segmentation using back propagation neural networks," *Cluster Computing*, vol. 22, no. S5, Article ID 12305, 2019.

[17] V. Gnanaprakash, P. T. Vanathi, and G. Suresh, "Defect detection in fabrics using back propagation neural networks," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 2s2, pp. 132–138, 2018.

[18] J. Shen, Z. Hu, R. P. Sharma et al., "Modeling height–diameter relationship for poplar plantations using combined-optimization multiple hidden layer back propagation neural network," *Forests*, vol. 11, no. 4, p. 442, 2020.

[19] V. Geetha, K. S. Aprameya, and D. M. Hinduja, "Dental caries diagnosis in digital radiographs using backpropagation neural network," *Health Information Science and Systems*, vol. 8, no. 1, pp. 8–1532, 2020.

[20] G. Roopali and T. Verma, "Tomato leaf disease detection using back propagation neural network," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 8, pp. 529–538, 2020.

[21] L. L. C. A. De, J. L. Lázaro-Galilea, A. Gardel Vicente et al., "Weak calibration of a visible light positioning system based on a position-sensitive detector: positioning error assessment," *Sensors*, vol. 21, no. 11, p. 3924, 2021.

[22] W. Cui, T. Wang, L. Cai, Y. Zhang, and J. Zhang, "Route optimization method of UAV autonomous transmission line detection based on BeiDou positioning service," *E3S Web of Conferences*, vol. 267, no. 1, Article ID 01055, 2021.

[23] O. Montenbruck, P. Steigenberger, L. Prange et al., "The multi-GNSS experiment (MGEX) of the international GNSS service (IGS) – achievements, prospects and challenges," *Advances in Space Research*, vol. 59, no. 7, pp. 1671–1697, 2017.

[24] J. Gilmore, M. Islam, J. Duncan, R. Natu, and R. Martinez-Duarte, "Assessing the importance of the root mean square (RMS) value of different waveforms to determine the strength of a dielectrophoresis trapping force," *Electrophoresis*, vol. 38, no. 20, pp. 2561–2564, 2017.

[25] Z. Bao, Y. Jiao, M. Chen, C. Zhang, and H. Li, "On-line anomaly detection for the measurement error of HVPT in the multi-bus structure," *Measurement*, vol. 170, no. 4, Article ID 108722, 2021.

Research Article

The Relationship between Physical Activity and Academic Achievement in Multimodal Environment Using Computational Analysis

Lingshu Li [1] **and Li Zhang** [2]

[1]*Department of Physical Education, Shanghai International Studies University, Shanghai 201620, China*
[2]*Department of Physical Education, China Pharmaceutical University, Nanjing, Jiangsu 211112, China*

Correspondence should be addressed to Li Zhang; zhanglicpu@126.com

Health has always been recognized as the imperative parameter to excel in any field whether professional or personal. People with sound health having regular habit of physical activities show more potential in their professional and personal lives than the people who do not participate in the physical activities. Many state-of-the-art studies exist in the literature where researchers have proved the significance of physical activities as supportive treatment for the existing ailments and in improving the overall health of human beings. Our research aims at accessing the correlation between physical activities carried out by the students and its impact on the success rate of academic achievements. The study is using computational techniques to investigate the relevance of physical activities on the academic achievements of middle school children. The study employs data mining techniques for processing the data. The computational methods are used in a multimodal environment where the surrounding parameters of the environment are considered before performing computational techniques on the subjects (participants in terms of sample for the study). In this cross-sectional study, we have considered the data on various physical activities such as aerobic fitness, running, playing, and participation in extracurricular activities. After collection of data in a real multimodal environment from middle school students, the data preprocessing is performed to handle the missing values. Then, the computational techniques are applied in a step-by-step approach using regression and the bootstrap methods to examine the data and predict the outcome. The correlation is assessed between academic achievements and physical activities. The outcome predicts that physical activities promote the success rate of academic achievements including extra-curricular activities.

1. Introduction

1.1. Background Study. As the focus on academic achievement has increased in the schools, physical activities have reduced. Health is the most important criterion, which impacts the human performance in all fields. Diet and physical activities play equally important role in the overall health of human beings [1]. Childhood obesity is a dominant threat to public health in the United States, which has been linked with physical health problems [2] and poorer academic achievement (AA) [3]. Children's health and AA can be improved through physical activity (PA) and aerobic fitness. To date, however, there is limited understanding of how indicators (PA, aerobic fitness, and attention) operate together jointly and separately to influence AA. Some inconsistencies in researches on the relationship between PA, aerobic fitness, and AA, which could be attributed to measure and methodological issues (e.g., measuring only PA or aerobic fitness and inadequate control of socio-economic variables); findings were either a positive or a negative association. The latter is rarely reported [4–6]. Therefore, the purpose of this study was to investigate the relation between PA and AA among middle school students.

1.2. Existing Literature. PA exerts the greatest effect on middle school and elementary students [7]. Among this population, active children generally show healthier cardiovascular features, aerobic fitness, and PA [8]. The effect of PA on cognitive ability was task dependent and was most evident in perceptual skills, followed by IQ [9, 10]. Previous studies have shown that PA has effects on enhancing angiogenesis, increasing oxygen saturation and glucose delivery, improving cerebral blood flow, and increasing neurotransmitter levels, which is conducive to children's cognitive function [11]. A study of 248 children (aged 8 to 11 years) indicated that PA was associated with VO2 peak, a measure of aerobic fitness, in both males and females [12]. However, there does not exist any study that correlate the relationship between PA, aerobic fitness, and AA. In separate models, the associations between PA and AA have been evaluated. PA, increasing physical effort, and social engagement can challenge executive functions and academic achievement. However, it fails to statistically make clear possible collinearity [13]. Some researchers have found that there is a lack of a relationship or a negative relationship between PA and AA. Moreover, the association between PA and aerobic fitness indicates a mediating effect of AA [14]. In reference [15], authors state that the influence of PA on AA may extend through the impact of PA on aerobic fitness. Aerobic fitness is an indicator of physical performance. Some studies have illustrated that aerobic fitness can not only promote hippocampal cerebral blood flow but also contribute to brain function and structure development [16]. How these measures (PA, aerobic fitness, and cognitive ability) interact and consequently influence AA is not yet clearly understood. Attention, as a major part of cognitive ability, has never been explored in the relationship between PA and AA. Without higher levels of attention and aerobic fitness, achieving and sustaining at least a moderate PA level might be precluded. Furthermore, there has been some evidence to prove that compared to physical activity, these physiological features are associated with academic achievement and cognitive function [17–20].

In reference [21], authors have carried out a survey to summarize the meta-analyses about the effect of PA on AA of school-going kids and teenagers. As per their findings, PA had shown null or small-to-medium positive effects on AA. Regular PA showed a medium-positive effect on AA, and rigorous PA did not reveal any benefits. In reference [22], authors conclude through their study that the PA and physical fitness are positively correlated with the AA of teenagers. They also conclude that the physical fitness criteria are closely related to the AA was cardio-respiratory fitness. In reference [23], authors identify two large gaps in the research carried out to correlate PA and AA: identifying the level and type of PA needed. The optimal type of PA to improve academic outcomes is also unknown. In reference [24] authors revealed that the overall effect of PA on AA was tilting towards positive; however, it was smaller than the total indirect impact through mediators. An indirect impact is the aggregate of the impact of self-confidence and depression, but the self-esteem factor was found to be the strongest mediator between PA and AA. Therefore, the present study aimed to evaluate the relationship between PA and AA and included attention, aerobic fitness, and other possible factors as mediators.

1.3. Major Highlights of the Proposed Study

(1) The cross-sectional study uses computational techniques to investigate the relevance of PA on the AA of middle school children and employs data mining techniques for collection of data and processing of data.

(2) The study is aimed to identify the effect of PA on the academic performance of the school going kids.

(3) The study also evaluates the correlation between MVPA (moderate-to-vigorous PA) and performance for mathematics scores of the participants

(4) The study uses multiple regression (MR) analysis and bootstrap analysis (BA) to conduct the experimental analysis.

The next section discusses about the proposed method.

2. Proposed Method

The proposed study uses regression analysis and BA to conduct the experimental analysis. The samples of the school going kids are collected to prepare the dataset. In order to evaluate the data, methods such as MRA and BA are applied as explained further in this section.

2.1. Multiple Regression Analysis (MRA).

It is an assessment of MR in statistics. It is an advancement of linear regression (LR). In statistics, LR is the used to forecast the value of one variable dependent on another variable. In MR, dependent variable depends on values of two or more variables. The MR analysis (MRA) estimates the information using regression.

2.2. Bootstrap Analysis (BA).

The bootstrap method is a resampling technique. It is used for the estimating the statistics on a population by sampling a dataset with replacement. It can be used to estimate mean or standard deviation (summary statistics).

This study uses MR analysis and BA to conduct the experiment of 176 seventh and eighth-grade students from USA and analyze the relationship between their mathematics scores and other measurements stated below. The following part will specifically introduce the research participants, measurement, study procedure, and data analysis process.

2.3. Participants or Subjects.

About 176 middle school students (from the 7th and 8th grade) from the school located in Salt Lake City, UT, USA are selected as subjects according to the cluster sampling technique [17–19]. Among them, 93 students were boys and 93 students were girls. An analytical cross-sectional analysis was performed. The study procedure was permitted by the institutional review board and the district research board. Written informed consent was taken from parents and students to conduct this research study.

2.4. Measurement. The specification of the measurement for each of our parameters has been explained as follows:

2.4.1. Physical Activity (PA). PA was estimated (maximum of 7 days per student) with a pedometer (NL-1000). Two PA measures were used as an independent variable: average number of steps for 1 week (steps) and average duration of moderate-to-vigorous PA (MVPA) for 1 week [16].

2.4.2. Body Mass Index (BMI). Height and weight were measured using a portable scale with an accuracy of 0.1 kg (BD-590; Tokyo, Japan) and a portable stadiometer with an accuracy of 0.01 m (Seca 213; Hanover, MD, USA), respectively. BMI was calculated (kg/m^2) and standardized; BMI z-score was obtained.

2.4.3. Aerobic Fitness. Using the PACER program, participants were asked to run as long as possible at a specified pace. They had to run back and forth across a 20 meter space at a fast pace every minute, with one point being added for every 20 meters [17].

2.4.4. Selective Attention. Selective attention was measured using the Stroop test. The Stroop test in the psychology is a measure of delay in reaction time. It indicates the flexibility of cognitive thinking The Stroop test was intended to measure selective attention and cognitive flexibility [18]. Our study followed the guideline according to previous studies: two types of paper cards were used to measure selective attention, and each card contained stimuli against a white background. On card A, every color's name was printed in black ink. On card B, serving as an interference card, color names did not match with ink color (e.g., black word printed in purple, green, or orange). Participants were asked to read out the words on card A as quickly as they can. For card B, they had to indicate the color of the printed word, not the word itself. The time taken to complete the task (all errors were corrected promptly without stopping the stopwatch) represented selective attention 1 (card A) and selective attention 2 (card B).

2.4.5. Sustained Attention and Alternating Attention. Sustained attention and alternating attention were measured through the trail making test (TMT). The original test was developed by Partington [19]. The task involves connecting 25 consecutive targets on a piece of paper or a computer screen in a way that resembles a "connect-the-dots" puzzle played by children. The test has two parts: the first part (TMTA) is the numbers test, where participants are asked to identify the numbers sequentially; the second part (TMTB) is a test, in which the child alternates between numbers and letters (1, A, 2, B, etc.) [22]. TMTA and TMTB measure sustained attention and alternating attention, respectively. Each participant practiced before the test and the test duration was determined. The examiner corrected errors immediately, without stopping the chronometer. The shorter the duration, the better is the sustained attention and alternating attention.

2.4.6. Academic Achievement (AA). AA was measured based on the students' mathematics score. The New York Test, which was designed by the New York State Education Department for testing students' mathematics level, was employed. There were 30 questions in the grade 7 edition and 27 in the grade 8 edition.

2.5. Procedure. Information about the study was shared with the students during their physical education (PE) classes. Consent forms were sent to the parents through the students; the parents had the chance to ask questions before starting the study. Child assent was completed during PE classes. During the first week, participants' height, weight, and aerobic fitness through a running test (PACER) were determined. During the second week, the PA test (i.e., steps and MVPA) was performed; measurement was obtained for 7 consecutive days. Starting from Monday, students started wearing a pedometer once they arrived at school and were required to wear the monitor all day until bedtime. The attention of the students was also evaluated. The participants were asked to complete the Stroop test and TMT; their mathematics scores were identified during their PE classes. BMI was calculated using height and weight measurements.

2.6. Data Analysis. All data were analyzed by SPSS 26.0 (IBM Corporation, Armonk, NY, USA). Variance analysis, least significant difference *t*-test, multivariate regression, and BA were conducted. These methods could help us understand the degree of the assumed positive correlation between PA and AA. We conducted the experiments with a sample size of 176 students, which was the average number in a typical public middle school in Utah; our study thus has sufficient statistical power. The sample size was also determined based on the number of schools willing to participate and their respective student membership size.

MR analysis was used to analyze the relationship between mathematics scores of grade 7 and 8 students and MVPA, number of steps, aerobic fitness, selective attention 1, selective attention 2, sustained attention, and alternating attention; the final regression equation was obtained. Regression analysis was performed by progressive stepwise regression and expressed as 95% confidence interval.

BA was used for analyzing mediating effects and was considered appropriate for the expected non-normality of the sampling distribution of the mediation effect [13]. Based on the direct and gradual effects between independent, dependent, and mediating variables, the mediating relationship was calculated for each variable.

3. Result Outcome

3.1. Descriptive Statistics of Grade 7 and 8 Students. Descriptive statistics are presented in Table 1. About 176 secondary school students enrolled in grades 7 and 8, 52.8%

Variable	Boys ($n = 93$)	Girls ($n = 83$)	Total
Height (cm)	163.68 (7.506)	157.84 (6.865)	160.93 (7.761)
Weight (kg)	62.47 (17.145)	62.15 (15.928)	62.32 (16.536)
BMI	23.1591 (5.50100)	24.7992 (5.34419)	23.9326 (5.47407)

Note. Values are shown as mean (standard deviation).

were female. Moreover, 52.8% were grade 7 students (48 boys and 45 girls) and 47.2% were grade 8 students (45 boys and 38 girls).

According to the World Health Organization (WHO) Growth Reference [17], pupils' BMI are divided into five categories: severely thin, thin, normal, overweight, and obese. The BMI classification of children aged 13 and 14 years was identified monthly; categories according to the WHO Growth Reference were used as standards (Table 2). The descriptive statistics of the BMI of grade 7 and 8 students are listed in Table 3.

The sample characteristics for middle school students in grades 7 and 8 are listed in Table 1.

The sample characteristics for middle school students in grades 7 and 8 are listed in Table 2.

The distribution of BMI in male and female students in grades 7 and 8 is listed in Table 3.

3.2. Multiple Linear Regression Analysis of PA and Potential Mediating Variables and Mathematics Scores.

Using experimental research, we have measured the indicators of each student, which are forming a categorical variable of BMI and multiple continuous variables. The forward stepwise regression method was used for the dependent variable Y: math and the independent variables $X1$: MVPA, $X2$: steps, $X3$: aerobic fitness, $X4$: selective attention 1, $X5$: selective attention 2, $X6$: sustained attention, $X7$: alternating attention, and $X8$: BMI. For BMI, dummy variables were recorded: $X8$: severe thinness, $X8$: thinness, $X8$: normal, $X8$: overweight, and $X8$: obesity. Regression analysis was performed to construct a MR model to test the interconnection between PA, potential mediating variables, and total academic scores.

MVPA and steps played a common role among the seventh and eighth-grade students and always had a positive effect on mathematics scores. In males, sustained attention had an effect on mathematics scores; selective attention 1, selective attention 2, alternating attention, and BMI were eliminated in the equation. The final equation of the multiple LR of the independent and dependent variables of males is written as follows:

$$\widehat{Y} = 5.839 + 1.302X1 + 0.004X2 + 0.657X6. \tag{1}$$

It could explain 47.7% of the total variation of the dependent variable (Table 4). Thus, MVPA steps and sustained attention could predict mathematics score to a certain extent. MVPA steps and sustained attention had positive effects on the improvement of mathematics scores in males. In females, the final equation of the multiple LR is expressed as follows:

$$\widehat{Y} = 6.327 + 1.610X1 + 0.005X2. \tag{2}$$

This could explain 60.1% of the total variation of the dependent variable (Table 5). Therefore, MVPA and steps could predict mathematics score to a certain extent, and MVPA and steps had positive effects on the improvement of mathematics scores in females.

The regression coefficient of the final model of boys' regression equation is listed in Table 4.

The regression coefficient of the final model of girls' regression equation is listed in Table 5.

3.3. Standardized Residual Analysis of the Regression Model.

The histogram has been plotted for the standardized residuals of the regression equations. The histograms of standardized residuals of the regression equations, from which standardized residuals showed normal distribution, conform to the hypothesis test condition (Figures 1 and 2). From the normal P-P graph of normalized residuals, the scattered point distribution of standardized residuals was close to the straight line, indicating that normalized residuals conform to normal distribution and residuals satisfy the conditions of multivariate LR (Figures 3 and 4).

Figures 5 and 6 show scatter plots of the mathematical scores and the standardized prediction value of regression, with the dependent variable on the x axis and ∗ZPRED on the y axis. The two variables were in a straight-line trend.

Therefore, step, MVPA, and sustained attention could predict mathematics score to a certain level. In males, the lower the average number of steps, the shorter the duration of MVPA, and the shorter the sustained attention time, the worse the mathematics score, which is contrary to the assumption. Moreover, the greater the average numbers of steps and longer the duration of MVPA, the better the total mathematics score. In females, sustained attention time did not influence the mathematics score. However, similar to males, the greater the average numbers of steps and longer the duration of MVPA, the better the total mathematics score.

Figure 1 shows the standardized residual histogram for the regression equation of boys.

Figure 2 shows the standardized residual histogram for the regression equation of boys.

Figure 3 shows regression-normalized residual P-P diagram of boys.

Figure 4 shows regression-normalized residual P-P diagram of girls.

We plotted the scatter plot of regression-standardized prediction values for boys and girls as shown in Figure 5. Figure 5 depicts scattered plot of regression-standardized prediction value (boys).

TABLE 2: Sample characteristics for middle school students in grades 7 and 8.

school Section	Severe thinness		Thinness		Normal		Overweight		Obesity	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
7	≤13.8	≤13.6	<14.9	<14.9	14.9–21.7	14.9–22.7	>21.7	>22.7	≥25.8	≥27.2
8	≤14.3	≤14.0	<15.5	<15.4	16.4–22.6	15.5–22.6	>22.6	>23.5	≥26.9	≥28.2

TABLE 3: Distribution of BMI in male and female students in grades 7 and 8.

Gender		N	Percentage
Boys	Severe thinness	0	0
	Thinness	1	1.1
	Normal	46	49.5
	Overweight	24	25.8
	Obesity	22	23.7
	Aggregate	93	100.0
Girls	Severe thinness	0	0
	Thinness	2	2.4
	Normal	31	37.3
	Overweight	23	27.7
	Obesity	27	32.5
	Aggregate	83	100.0

Figure 6 shows the scatter plot of the regression-standardized prediction value (girls).

3.4. Test for Mediating Effect. We have carried out the BA to find out the mediating effect of MVPA on the mathematical achievement of students and the same has been explained as below.

3.4.1. Analysis of the Mediating Effect of MVPA on the Mathematical Achievement of Students. Table 6 lists the result of the mediating effect analysis by the bootstrap method, where X is the independent variable (MVPA) and Y is the dependent variable (mathematics achievement). The direct effect of MVPA on mathematics achievement in males and females was significant (95% confidence interval: 1.3906–3.4871 and 1.5650–5.4291, respectively). The indirect effect of each mediating variable was not significant; thus, MVPA has no mediating effect on mathematics achievement.

3.4.2. Analysis of the Mediating Effect of Step on Mathematical Achievement. Table 7 presents the result of the mediating effect analysis by the bootstrap method, where X is the independent variable (step) and Y is the dependent variable (mathematics achievement). The direct effect of step on mathematics achievement of male and female students is significant (95% confidence interval: 0.0034–0.0059 and 0.0039–0.0065, respectively). However, the indirect effect of each mediating variable is not significant; thus, step has no mediating effect on mathematical achievement.

Analysis of the mediating effect of MVPA on mathematical achievements is listed in Table 6.

Analysis of the intermediary effect of step on mathematical achievement is listed in Table 7.

4. Discussion

We have used empirical research to explore the correlation between PA and AA and it has been found that PA could affect students' mathematics performance to a certain extent. In males, the multiple LR equation has shown that mathematics scores were related to MVPA steps and sustained attention. In females, the multiple LR equation has shown that mathematics scores were related to MVPA and steps. To summarize, PA has a direct effect on AA and there are no mediating effects between PA and three kinds of attention, namely, selective attention, sustained attention, and alternating attention.

One possible explanation of our findings is that individuals with more agile bodies have higher flexibility and achieve greater number of steps. Mathematics requires higher mental function and coordination between the hands and brain. The ability to solve mathematics problems requires using the brain and the using a paper, pen, and tools for calculations [21].

Moreover, MVPA has an effect on mathematics performance because proper PA could improve blood circulation, thereby fully mobilizing the functions of various organs of the body, including the brain [22]. PA has a positive effect on the central nervous system, such as enhanced flexibility of the central nervous system and coordination with various systems, improved ability to respond and keep the balance of the interaction between excitatory and inhibition processes, and improved regulation of the central nervous system and muscle activity of the body [23]. These effects could in turn make individuals more energetic, thereby improving learning efficiency [4]. Moreover, PA could help stabilize students' emotions, thus enabling students withstand greater academic pressure. A team of experts who studied the significance of PA in school-age children proved that active children and adolescents are more encouraged to improve their AA before, during, and after school because exercising and fitness contributed to the development of children's brain structure, function, and cognitive ability [15]. Moreover, middle school is the golden age of physical and intellectual development. For students, the burden of learning is not considered extremely heavy because they have much spare time. Moreover, the longer one engages in PA, the more exercise his or her body gets and their physical condition improves, which in turn provides sufficient energy and a strong body for learning. Exercise could also enrich students' extracurricular activities, stimulates the release of a variety of chemicals such as dopamine and endorphin, thus enhancing memory and intellectual development and promoting happiness. An optimistic attitude or a positive mental outlook is also conducive for improving learning ability.

TABLE 4: Regression coefficient of the final model of boys' regression equation.

Coefficient[a]						
Model		Unnormalized coefficient		Normalized coefficient	T	Significance
		B	Standard error	Beta		
Boys	Constant	5.839	5.855		0.997	0.321
	Step	0.004	0.001	0.535	6.293	0.0001
	MVPA	1.302	0.435	0.254	2.993	0.004
	Sustained attention	0.657	0.316	0.160	2.078	0.041

Note. [a]Dependent variable: math.

TABLE 5: Regression coefficient of the final model of girls' regression equation.

Coefficient[a]						
Model		Unnormalized coefficient		Normalized coefficient	t	Significance
		B	Standard error	Beta		
Girls	Constant	6.327	4.504		1.405	0.164
	Step	0.005	0.001	0.594	7.649	0.0001
	MVPA	1.610	0.406	0.308	3.963	0.0001

Note. [a]Dependent variable: math.
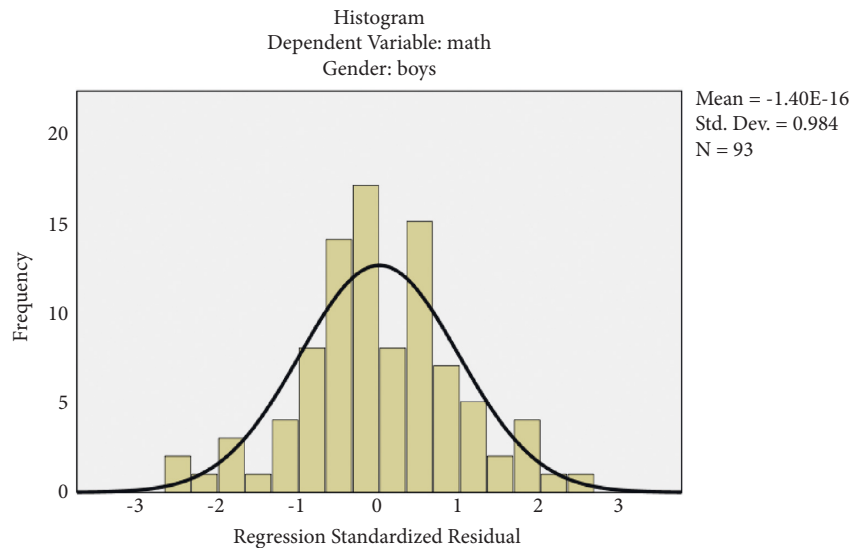
FIGURE 1: Standardized residual histogram for the regression equation of boys. The regression-standardized residuals show normal distribution, which means that the data meet the requirements of the regression equation (same is applicable for Figures 2 to 6).

BMI and aerobic fitness do not have any effect on mathematics scores because these obesity indicators have little effect on the students' willingness to exercise and exercise preferences, and the ability to solve mathematical problems is more related to the understanding and application of mathematical knowledge [16]. In this study, sustained attention levels did not show a mediating effect, and the analysis of sex differences showed contradicting results. Sustained attention could only be effective in predicting mathematical performance (direct effect and indirect effect) under certain conditions [21]. Moreover, sex differences in terms of persistent attention exist, which may explain the results of this study to some extent. Moreover, the absence of the mediating effect of sustained attention could be attributed to the fact that sustained attention is extremely limited. Even in the most ideal interference-free environment, an individual's attention span could only last for 20–22 min. The time required for measuring AA is much longer than the duration of the individual's attention; thus, the effect of sustained attention is extremely limited.
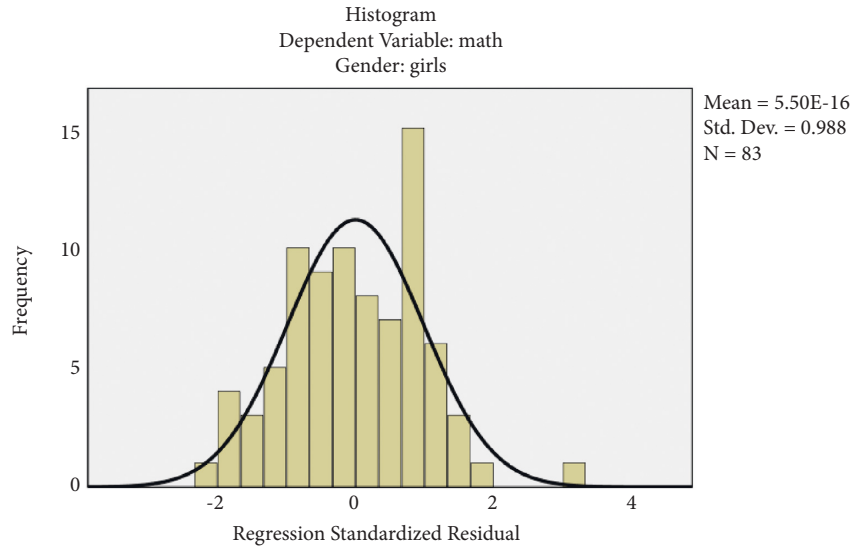
Figure 2: Standardized residual histogram for the regression equation of girls. The regression-standardized residuals show normal distribution.
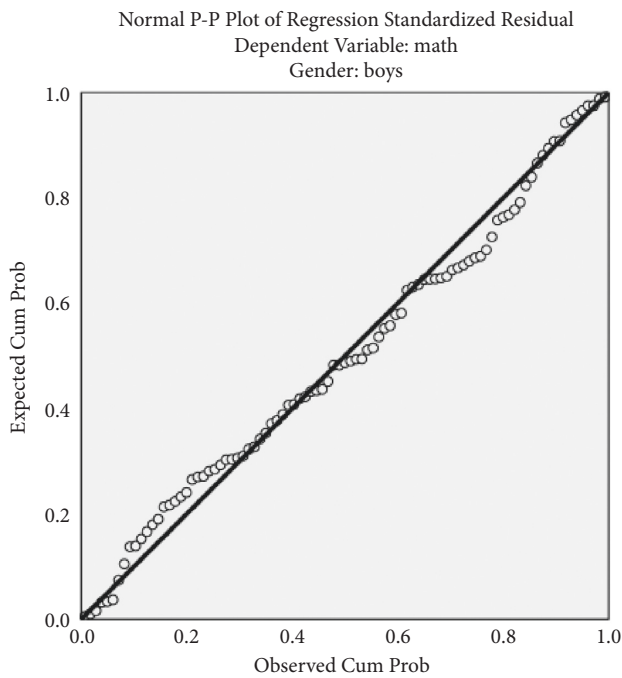
Figure 3: Regression-normalized residual P-P diagram of boys. The scattered point distribution of standardized residuals is close to the straight line.
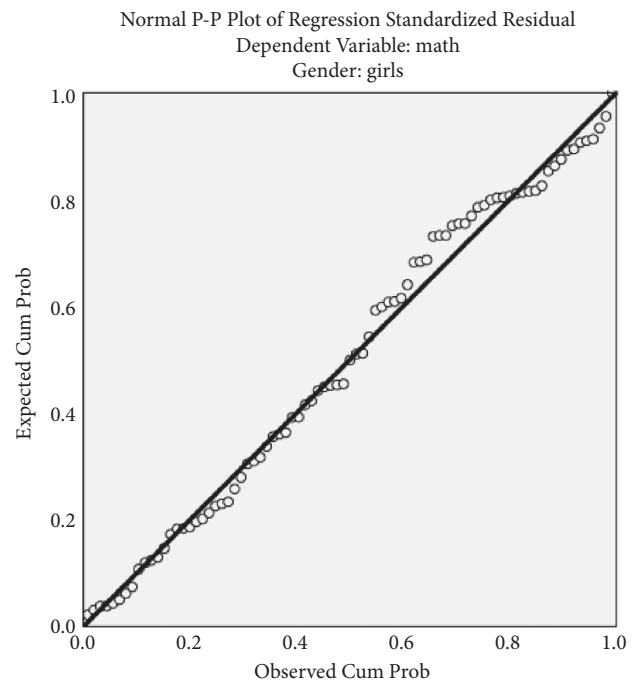
Figure 4: Regression-normalized residual P-P diagram of girls. The scattered point distribution of standardized residuals is close to the straight line.

Therefore, the level of sustained attention did not produce a prominent effect in this study.

Furthermore, alternating attention was a form of selective switching attention. In our study, the effect of divergent attention on the student's learning process was not significant. For students, learning for or taking exams is what they usually have to face. The diversion of attention and demand for attention distribution are not high. The time for a person to complete a color test could only reflect his/her sensitivity to color. Thus, Stroop test results are not associated with mathematics scores. For selective attention, selective attention could be correlated with AA to a certain extent. However, AA could not be predicted with selective attention, which is consistent with the finding of our study. Moreover, although selective attention has no direct influence on AA, it has an indirect effect on working memory when it is needed, which is used as a mediating variable. Thus, AA could be influenced by the combination of
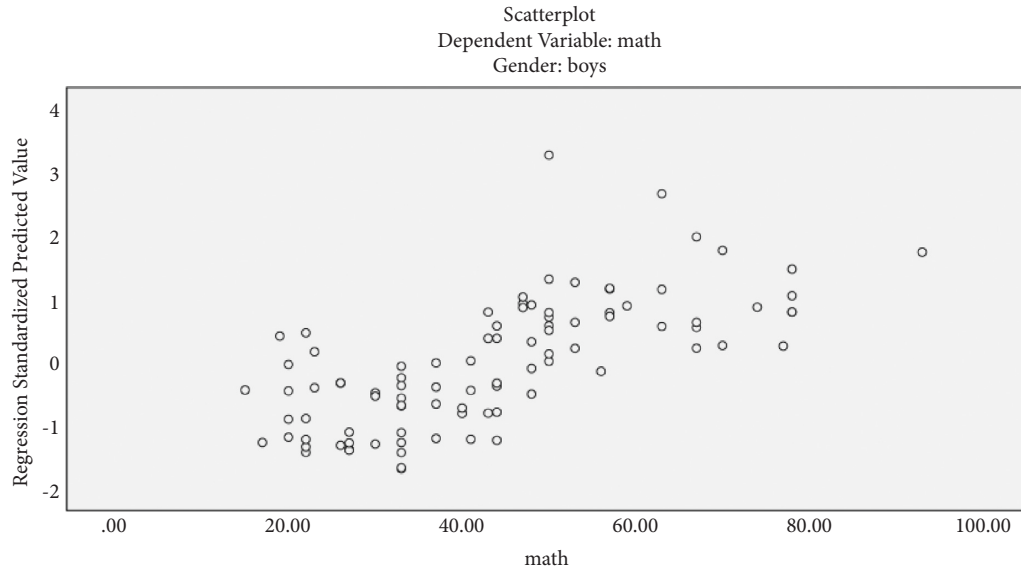
FIGURE 5: Scatter plot of regression-standardized prediction value (boys). Dependent variable is placed on the *x* axis and * ZPRED is the *y*-axis variable. The two variables are in a straight-line trend.
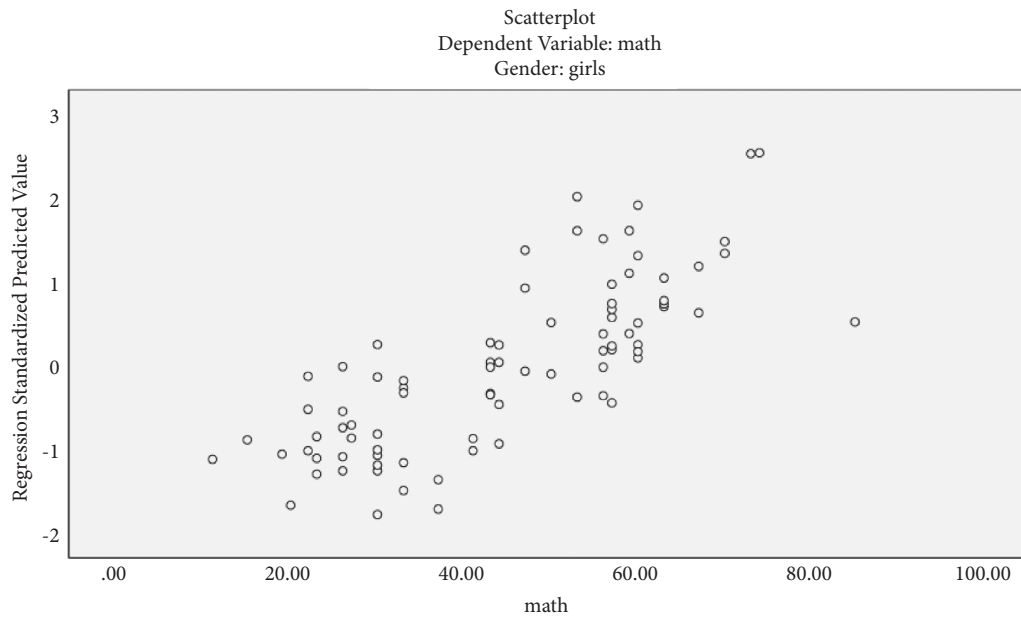
FIGURE 6: Scatter plot of the regression-standardized prediction value (girls). Dependent variable is placed on the *x* axis and * ZPRED is the *y*-axis variable. The two variables are in a straight-line trend.

sustained attention, alternating attention, and selective attention as mediating variables of PA. However, based on the BA, we found no mediating effects between PA and the three kinds of attention. Hence, whether selective attention, sustained attention, and alternating attention have direct or indirect effects on PA and AA requires further investigation, and future studies on other mediating variables between the three kinds of attention and AA are warranted. There are some limitations to this study. First, the sample in our study involved only one school. This study would need to be replicated in other areas in the USA and around the USA for greater transferability. Moreover, more factors regarding PA and psychology could be explored and considered for a more accurate statistical analysis.

TABLE 6: Analysis of the mediating effect of MVPA on mathematical achievements.

Direct effect of X on Y			Indirect effect (s) of X on Y								
					Boys				Girls		
	Boys	Girls		Effect	BootSE	BootLLCI	BootULCI	Effect	BootSE	BootLLCI	BootULCI
Effect	2.4389	3.4970	Total	0.0069	0.2879	−0.5624	0.5947	−0.5922	0.7934	−2.3109	0.8381
se	.5273	0.9701	Aerobic fitness	0.0811	0.2346	−0.4045	0.5437	−0.4304	0.7334	−2.0494	0.8330
t	4.6252	3.6050	Selective attention 1	−0.0510	0.0985	−0.2617	0.1398	−0.0258	0.0747	−0.1861	0.1285
p	.0001	0.0006	Selective attention 2	0.0224	0.1208	−0.1810	0.3193	−0.1083	0.1589	−0.4627	0.1781
LLCI	1.3906	1.5650	Sustained attention	−0.0351	0.1058	−0.2668	0.1783	−0.0276	0.2042	−0.4211	0.4213
ULCI	3.4871	5.4291	Alternating attention	−0.0106	0.0746	−0.1884	0.1379	0.0001	0.1143	−0.2600	2378

TABLE 7: Analysis of the intermediary effect of step on mathematical achievement.

Direct effect of X on Y			Indirect effect (s) of X on Y								
					Boys				Girls		
	Boys	Girls		Effect	BootSE	BootLLCI	BootULCI	Effect	BootSE	BootLLCI	BootULCI
Effect	0.0046	0.0052	Total	0.0001	0.0003	−0.0006	0.0007	0.0005	0.0003	−0.0001	0.0011
se	0.0006	0.0007	Aerobic fitness	0.0001	0.0002	−0.0004	0.0005	0.0006	0.0003	0.0001	0.0011
t	7.2668	7.9861	Selective attention 1	0.0001	0.0001	−0.0001	0.0004	0.0001	0.0001	−0.0003	0.0001
p	0.0001	0.0001	Selective attention 2	0.0001	0.0002	−0.0004	0.0003	0.0001	0.0001	−0.0003	0.0002
LLCI	0.0034	0.0039	Sustained attention	−0.0001	0.0001	−0.0003	0.0001	0.0001	0.0001	−0.0002	0.0002
ULCI	0.0059	0.0065	Taril_M2	0.0001	0.0001	−0.0002	0.0002	−0.0001	0.0002	−0.0004	0.0003

5. Conclusions

This research study has carried out the empirical research on the identified population. The findings of the paper state that PA has a direct effect on AA. There are no mediating effects between PA and three kinds of attention, namely, selective attention, sustained attention, and alternating attention. No mediating effects are observed between PA and the three kinds of attention, namely, selective attention, sustained attention, and alternating attention. In males, the multiple LR equation has shown that mathematics scores were related to MVPA and sustained attention. In females, the multiple LR equation showed that mathematics scores were related to MVPA and steps. From the above findings, it can be concluded that the short time spent on physical activities during PE classes and sports related activities impacts positively on the academic performances of middle school students. It is found that there is a strong correlation between the physical activities and overall performance of the participant subjects. This is due to the hormones, which are released by the body during physical activities that relieve stress and increase concentration levels. This has resulted into the greater success in terms of mathematics scores of the middle school children. This study is applicable to all the age groups but in this research study, we are focusing on the school going kids and the impact of physical activities on the academic performance is remarkable as per the research study. The study of biomarkers also show that the relevance of physical activities on the overall performance of the students. The computational techniques like regression were used in predicting the performance on the basis of input activity data of the participants. The future study will concentrate on analyzing the impact of physical activities on the participants by considering the most relevant biomarkers of health analysis. We will introduce better computational techniques with different datasets to draw conclusive remarks to analyze the impact of physical activities on the academics. Furthermore, we will explore more factors related to PA and psychology in the future.

Data Availability

The data used to support the findings of the study can be obtained from the first author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] S. Judge and L. Jahns, "Association of overweight with academic performance and social and behavioral problems: an update from the early childhood longitudinal study," *Journal of School Health*, vol. 77, no. 10, pp. 672–678, 2007.

[2] M. Kaur and S. Kadam, "Bio-inspired workflow scheduling on HPC platforms," *Tehnički Glasnik*, vol. 15, no. 1, pp. 60–68, 2021.

[3] E. K. Howie and R. R. Pate, "Physical activity and academic achievement in children: a historical perspective," *Journal of Sport and Health Science*, vol. 1, no. 3, pp. 160–169, 2012.

[4] M. Kaur, "Elitist multi-objective bacterial foraging evolutionary algorithm for multi-criteria based grid scheduling problem," in *Proceedings of the 2016 Int. Conf. on Internet of Things and App. (IOTA)*, pp. 431–436, Pune, India, January 2016.

[5] L. Chaddock, M. B. Pontifex, C. H. Hillman, and A. F. Kramer, "A review of the relation of aerobic fitness and physical activity to brain structure and function in children," *Journal of the International Neuropsychological Society*, vol. 17, no. 6, pp. 975–985, 2011.

[6] W. Zhang and M. Kaur, "A novel QACS automatic extraction algorithm for extracting information in blockchain-based systems," *IETE Journal of Research*, vol. 2022, Article ID 2030252, 13 pages, 2022.

[7] N. M. Mule, D. D. Patil, and M. Kaur, "A comprehensive survey on investigation techniques of exhaled breath (EB) for diagnosis of diseases in human body," *Informatics in Medicine Unlocked*, vol. 26, Article ID 100715, 2021.

[8] M. Kaur, S. R. Sakhare, K. Wanjale, and F. Akter, "Early stroke prediction methods for prevention of strokes," *Behavioural Neurology*, vol. 2022, Article ID 7725597, 2022.

[9] M. Kaur, A. Jadhav, and F. Akter, "Resource selection from edge-cloud for IIoT and blockchain-based applications in industry 4.0/5.0," *Security and Communication Networks*, vol. 2022, p. 9314052, 2022.

[10] C. Boreham and C. Riddoch, "The physical activity, fitness and health of children," *Journal of Sports Sciences*, vol. 19, no. 12, pp. 915–929, 2001.

[11] G. L. Curtis, M. Chughtai, A. Khlopas et al., "Impact of physical activity in cardiovascular and musculoskeletal health: can motion be medicine?" *Journal of Clinical Medicine Research*, vol. 9, no. 5, pp. 375–381, 2017.

[12] N. A. Khan and C. H. Hillman, "The relation of childhood physical activity and aerobic fitness to brain function and cognition: a review," *Pediatric Exercise Science*, vol. 26, no. 2, pp. 138–146, 2014.

[13] T. Tanha, P. Wollmer, A. Fedorowski, O. Thorsson, M. K. Karlsson, and M. Dencker, "Correlation between physical activity, aerobic fitness and body fat against autonomic function profile in children," *Clinical Autonomic Research*, vol. 26, no. 3, pp. 197–203, 2016.

[14] J.-S. Ryu, H. R. Chung, B. M. Meador, Y. Seo, and K.-O. Kim, "The associations between physical fitness, complex vs simple movement, and academic achievement in a cohort of fourth graders," *International Journal of Environmental Research and Public Health*, vol. 18, no. 5, p. 2293, 2021.

[15] P. D. Tomporowski, C. L. Davis, P. H. Miller, and J. A. Naglieri, "Exercise and children's intelligence, cognition, and academic achievement," *Educational Psychology Review*, vol. 20, no. 2, pp. 111–131, 2008.

[16] S. P. Harvey, K. Lambourne, J. L. Greene, C. A. Gibson, J. Lee, and J. E. Donnelly, "The effects of physical activity on learning behaviors in elementary school children: a randomized controlled trial," *Contemporary School Psychology*, vol. 22, no. 3, pp. 303–312, 2018.

[17] A. Diamond and D. S. Ling, "Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not," *Developmental Cognitive Neuroscience*, vol. 18, pp. 34–48, 2016.

[18] C. N. Rasberry, S. M. Lee, L. Robin et al., "The association between school-based physical activity, including physical education, and academic performance: a systematic review of the literature," *Preventive Medicine*, vol. 52, pp. S10–S20, 2011.

[19] D. M. Pindus, E. S. Drollette, M. R. Scudder et al., "Moderate-to-vigorous physical activity, indices of cognitive control, and academic achievement in preadolescents," *The Journal of Pediatrics*, vol. 173, pp. 136–142, 2016.

[20] T. Ishihara, N. Morita, T. Nakajima, K. Okita, K. Yamatsu, and M. Sagawa, "Direct and indirect relationships of physical fitness, weight status, and learning duration to academic performance in Japanese schoolchildren," *European Journal of Sport Science*, vol. 18, no. 2, pp. 286–294, 2018.

[21] A. Barbosa, S. Whiting, P. Simmonds, R. Scotini Moreno, R. Mendes, and J. Breda, "Physical activity and academic achievement: an umbrella review," *International Journal of Environmental Research and Public Health*, vol. 17, no. 16, p. 5972, 2020.

[22] C. C. Rodriguez, E. M. D. Camargo, C. R. R. Añez, and R. S. Reis, "Physical activity, physical fitness and academic achievement in adolescents: a systematic review," *Revista Brasileira de Medicina do Esporte*, vol. 26, p. 2020, 2020.

[23] E. K. Howie and R. R. Pate, "Physical activity and academic achievement in children: a historical perspective," *Journal of Sport and Health Science*, vol. 1, no. 3, pp. 160–169, 2012.

[24] S. Kayani, T. Kiyani, J. Wang, M. L. Z. Sánchez, S. Kayani, and H. Qurban, "Physical activity and academic performance: the mediating effect of self-esteem and depression," *Sustainability*, vol. 10, p. 3633, 2018.

Research Article

Using Real-Time fMRI Neurofeedback to Modulate M1-Cerebellum Connectivity

Yahia Madkhali [ID],[1] **Salim Al-Wasity,[2] Norah Aldehmi,[2] and Frank Pollick[3]**

[1]*Faculty of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia*
[2]*College of Medical, Veterinary and Life Sciences (MVLS), University of Glasgow, Glasgow, UK*
[3]*School of Psychology, University of Glasgow, Glasgow, UK*

Correspondence should be addressed to Yahia Madkhali; y.madkhali.1@research.gla.ac.uk

Objective. The potential of neurofeedback to alter the M1-cerebellum connectivity was explored using motor imagery-based rt-fMRI. These regions were chosen due to their importance in motor performance and motor rehabilitation. *Methods*. Four right-handed individuals were recruited to examine the potential to change the M1-cerebellum neurofeedback link. The University of Glasgow Cognitive Neuroimaging Centre used a 3T MRI scanner from January 2019 to January 2020 to conduct this prospective study. Everyone participated in each fMRI session, which included six NF training runs. Participants were instructed to imagine complicated hand motions during the NF training to raise a thermometer bar's height. To contrast the correlation coefficients between the initial and last NF runs, a *t*-test was performed post hoc. *Results*. The neurofeedback connection between M1 and the cerebellum was strengthened in each participant. Motor imagery strategy was a significant task in training M1-cerebellum connectivity as participants used it successfully to enhance the activation level between these regions during M1-cerebellum modulation using real-time fMRI. The *t*-test and linear regression, on the other hand, showed this increase to be insignificant. *Conclusion*. A novel technique to manipulate M1-cerebellum connectivity was discovered using real-time fMRI NF. This study showed that each participant's neurofeedback connectivity between M1 and cerebellum was enhanced. This increase, on the other hand, was insignificant statistically. The results showed that the connectivity between both areas increased positively. Through the integration of fMRI and neurofeedback, M1-cerebellum connectivity can be positively affected.

1. Introduction

Cerebellar functions have received a lot of attention over the previous decades. The timing and accuracy of skilled motions rely on the cerebellum [1]. Functional neuroimaging in humans has revealed that motor learning and performance are linked to the cerebellum. The right motor cortex is linked to the left cerebellar hemisphere. If the left side of the cerebellum is damaged, the right side will be impacted. The cerebellum activation included a range of sensory and motor activities, according to previous functional magnetic resonance imaging (fMRI) studies. The cerebellum helps with several cognitive processes, such as motor imagery [2]. Several functional imaging studies have been conducted on

the cerebellum. The first functional study of the cerebellum was conducted by Fox and colleagues [3]. They discovered that voluntary finger movement-activated bilateral lobule V. During a finger-tapping exercise, Kuhtz-Buschbeck, and colleagues [2] saw activity in the anterior cerebellum using fMRI.

The cerebellar regions that execute motor activities are different from those that execute cognitive and non-motor activities, according to Stoodley and Schmahmann [4]. Motor functions are controlled by the anterior cerebellum, while the posterior cerebellum controls cognitive functions. Since it aids in the pathophysiology of motor disorders linked to the cerebellum, understanding how cortical regions connect is critical. The cerebellum and the primary motor

cortex (M1) were indirectly linked, according to Hoover and Strick [5], via MRI investigations. Daskalakis et al. [6], however, have proposed an indirect cross-link via the inferior parietal lobe; thus a direct connection to M1 is unlikely. Additionally, the cerebellum may be less directly connected with other cortical regions than that of M1, according to Manto et al. [7]. Many research such as functional connectivity in autism [8] and Parkinson's disease [9] have been conducted to investigate the link between the motor cortex and cerebellum. Ramos and colleagues [10] also utilized resting-state fMRI to investigate alterations in intrinsic cortico-cerebellar functional linkages in individuals with typical development. As a result of their research, Mostofsky and his colleagues examined youngsters with typical development (TD) and high-functioning autism (HFA) [11].

As a consequence of this investigation, the contralateral thalamus, ipsilateral cerebellum, contralateral main sensory cortex, and SMA were all engaged in motor execution. The contralateral thalamus, ipsilateral cerebellum, contralateral main sensory cortex, and SMA were all engaged in motor execution. The study found that the level of engagement was specific to the movement being executed, for example, when executing a hand gesture with the left hand, these regions were more active than when executing the same hand gesture with the right hand. This suggests that these regions are specifically activated during certain tasks or movements and may be important for their specific function. Moreover, these processes were demonstrated to be differentially engaged between TD children and ASD individuals. As a result of their investigation, Ramos et [10] suggested: "It might also help in targeting neuroimaging investigations that new treatment strategies eventually fade out as older or less potent drugs are substituted for them." The authors further explored the specific areas of distributed abnormalities that comprised this fMRI activity in each situation.

Individuals may internally replicate actions or movements using Motor Imagery (MI) technique. Individuals may imagine about achieving motor things, but no physical outcomes are allowed in MI [11]. MI comes in two flavors: visual imagination (VI) and kinaesthetic imagination (KI) [11]. MI has a significant function since it is often utilized in motor learning exercises [12]. It may be beneficial to athletes, skill development, and rehabilitation since it may be used to improve motor performance over periods [13]. When physical movement, such as clenching or tapping, was imagined by participants using MI, the cerebellum was activated [10]. According to a study by Blefari et al. [12], even though cerebellar activity was high during motor execution, it decreased by 30% during MI. Neurofeedback is using visual or auditory stimuli to assist in self-regulation in a person [12]. Real-time fMRI (rt-fMRI) systems have been used for neurofeedback because of their speed and capacity to supply signal feedback from the brain activity of the subcortical brain framework. Neurofeedback has been shown to be an effective treatment for a wide range of disorders, including depression and anxiety. One potential mechanism by which neurofeedback may work is by

Table 1: Types of validity tests: factorial, construct, and concurrent.

	M1-cerebellum (mean)
Age (years)	36
Handedness	92.5
MI vividness (third-person perspective)	22.75
MI vividness (first-person perspective)	21.5

modulating brain function. Previous studies have shown that motor imagery-based rt-fMRI connectivity is associated with cerebellum function. The current study aim is to determine whether neurofeedback using motor imagery-based rt-fMRI connectivity is able to modulate M1-cerebellum connectivity.

The current research aimed to see whether people's brain feedback connections between the primary motor cortex (M1) and cerebellum might be changed. These adjustments may provide insight into healthy people's link between the cerebellum and motor cortex. The M1-cerebellum coupling has never been studied in conjunction with rt-fMRI modulation. Modifying the M1-cerebellum connectivity may help in illness therapy such as Parkinson's disease and autism by improving motor functions [11, 16].

2. Methods

2.1. Participants. This study included four individuals (three men and one woman) with normal or corrected vision. According to the participants, all the material and photos on the screen were viewable. According to the Edinburgh Handedness Inventory [17], all participants were right-handed. English was a breeze for all participants, whether they spoke or wrote it. Participants' average age was 36 years old. Each participant completed the Vividness of Movement Imagery Questionnaire-2 (VMIQ-2) to assess their capacity to perform the motor imagination tasks. The VMIQ-2[18, 19] is a 12-question exam that measures an individual's ability to generate new motions. There are two types of motor imagery: internal (first person) and external (third person) visual imagery. As seen in Table 1, the VMIQ-2 has passed all three types of validity tests: factorial, construct, and concurrent.

This study received ethical approval from the College of Science and Engineering ethics committees at the University of Glasgow. Each participant approved the experiment. Before undergoing the MRI scanning, each participant received a thorough explanation of the experiment's complete technique; all participants completed an MRI safety checklist.

2.2. Imaging Parameters and Rt-fMRI Neurofeedback Platform. Between January 2019 and January 2020, the University of Glasgow Centre for Cognitive Neuroimaging (CCNi) conducted this prospective study using a 3T Siemens Tim Trio MRI scanner with a 32-channel head coil. The Echo Planar Imaging (EPI) technique (TR = 2000 ms, 0.3 mm gap, 32 axial slices, TE = 30 ms) was used to acquire $T2$-weighted functional scans. All participants in this study underwent a high-resolution anatomical scan ($T1$ weighted image), six
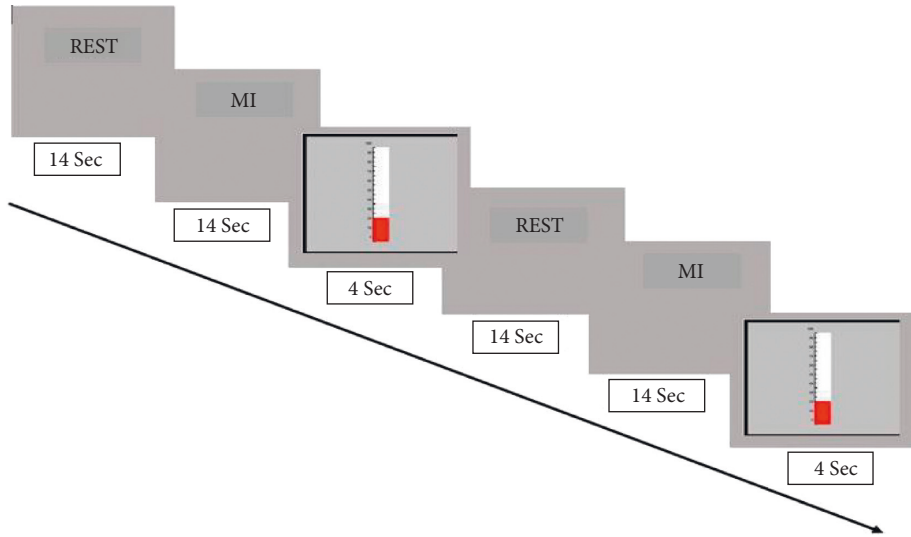
FIGURE 1: The NF training paradigm is constructed using fMRI. When the thermometer bar was shown for 4 seconds, a run lasted 366 seconds and included twelve 14 s fixations (rest) blocks and eleven 14 s NF blocks.

NF runs, and a functional localizer run. The scanner provided Turbo-BrainVoyager (TVB) software with useful data over a network connection during NF runs. Preprocessing of the functional data took place in real-time. The correlation between the two regions was calculated using custom MATLAB code, which was then displayed as a thermometer bar on the computer. During their NF training, the participants were encouraged to imagine complex hand motions to raise the height of the thermometer bar. For the motor imagery task, there were no particular hand motions recommended. Participants were able to test different mental strategies and actions to raise their thermometer. The height of the thermometer bar was determined using the Pearson correlation coefficient. The thermometer bar level was updated and displayed after each NF block using an intermittent feedback paradigm. During the localizer run, 7 fixation blocks (16 s) intermingled with 6 bimanual hand clenches (the 30 s). Participants were supposed to count letters or numbers when "REST" appeared and clench their fists when "Clench" occurred during functional scanning (the localizer). The block labeled "Clench" lasts 30 seconds, whereas the block labeled "Rest" lasts 16. The available data were processed and analyzed online using BrainVoyager. Each participant's original space was used to generate the ROIs.

2.3. Rt-fMRI Neurofeedback. Every participant completed 366-second NF training runs. Every NF training run included 32 s blocks. Each cycle consisted of a 14-second break (twice repeated throughout each cycle since each cycle began and ended with a rest), an 11-second MI exercise, and four seconds of feedback signal (Figure 1). During the blocks, participants were told to perform imagination of complicated motions for 14 seconds to raise the activation level between the M1-cerebellum connection. After each NF block, the thermometer bar displayed M1-cerebellum integration for 4 seconds. Participants were instructed to count

numbers or letters for 14 seconds to keep track of their baseline activity [17].

2.4. Online Data Analysis. Turbo-BrainVoyager software and MATLAB were used to analyze and present rt-fMRI data. A network link was used to transfer the data to Turbo-BrainVoyager software. In real-time, the available data were preprocessed. To update and display the feedback signal, which included an intermittent thermometer, the following equation was used:

$$\textbf{bar height} = \text{scale} * (\text{Corr} + 1). \tag{1}$$

To improve the visibility of the thermometer bar, the scale is an arbitrary scale value.

$$\frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}}. \tag{2}$$

The mean activation of $M1$ during a MI block is represented by x, which is a 14-s time course of M1, and the mean activation of the cerebellum is represented by y.

2.5. Full Brain Analysis. BrainVoyager QX 2.8.4 (Brain Innovation, Maastricht, The Netherlands) was used to manipulate the available data, which was then imported into MATLAB. Before being normalized to Talairach space, inhomogeneity was corrected in the anatomical $T1$ weighted image. Moreover, all volumes were spatially aligned to the localizer run's first volume to eliminate head motion, and the functional data were slice-time corrected. After that, Brain Voyager was used to align the fMRI data to $T1$ weighted image using anatomical landmark points if necessary. A rigid body transform and scaling were used to convert the data to Talairach space. Nonlinear drifts in the time series were eliminated using a temporal high-pass filter (2 cycles) applied to linear drifts. A Gaussian kernel with a 6-mm full width at
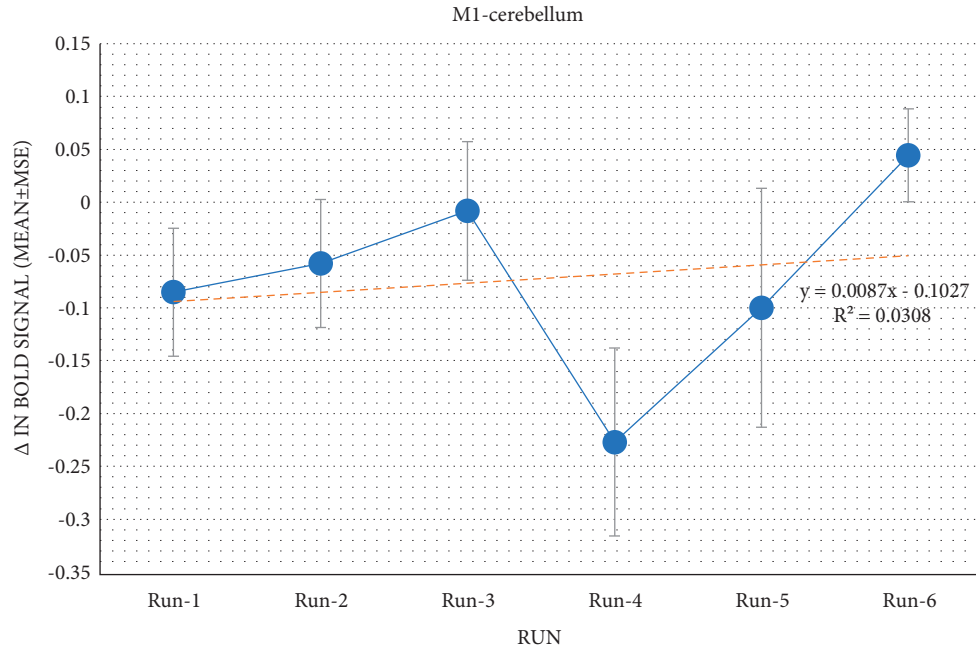
FIGURE 2: When comparing the first and last runs, this statistic shows the average BOLD signal change in $M1$-cerebellum connectivity. The standard deviation of the mean represented the error bars.

half maximum was used to smooth the functional data. To analyze preprocessed functional data with two predictors (fist clenching and rest) and six head motion parameters, a General Linear Model (GLM) was utilized for each subject [18]. A second-level random-effect analysis general linear model (RFX-GLM) was used to test group data. The generated statistical maps were thresholded and corrected for multiple comparisons using cluster-level thresholding [19]. The blocks of the NF were then contrasted to the baseline ($p < 0.01$ corrected for cluster-level thresholding) in all six NF runs.

2.6. Statistical Analysis. A t-test was used as a post hoc test to examine the correlation between the first and last runs. Linear regression was also performed using the average connectivity values of each neurofeedback run in order to calculate the upregulation over NF runs.

3. Results

The participants' cerebellar and M1 activation was observed during the localizer, owing to hand clenching. Each participant completed all six runs successfully. Participants seemed to upregulate activation between M1 and the cerebellum before dropping at run 4. Figure 2 shows an upwards trend in activation levels across runs.

Nonetheless, participants showed a trend for successfully upregulating based on the linear regression results (Intercept coefficient = −0.103, $R^2 = 0.272$ and p value = 0.0739). These findings suggested effective modulation due to differences in activation levels from run 1 to run 6.

According to the paired t-test, the observed increase in correlation was not statistically significant ($p = 0.287$). As result, four participants were not able to upregulated the

neurofeedback connectivity between M1 and the cerebellum using real-time fMRI.

In addition, participants thought playing music, knocking on doors, or lifting weights might help them raise their thermometer bar as shown in Figure 3.

3.1. Whole Brain Analysis. To see if any brain areas were activated during NF-directed motor imagery, a comprehensive RFX-GLM analysis ($p < 0.01$ corrected) was performed. Whole brain analysis shows that activations were seen in the cerebellum and M1 where the feedback was received. Also, bilateral parietal lobes, bilateral insula, and caudate body were all activated. Table 2 and Figure 4 list and illustrate the activated areas during this real-time fMRI experiment.

4. Discussion

In a single session of motor imagery of complex body motions with intermittent feedback signals (represented as the height of a thermometer bar) using real-time fMRI, we discovered preliminary evidence that healthy people might learn to increase their activity level between $M1$ and cerebellum connectivity. To our knowledge, real-time fMRI has never been used to investigate the relationship between cortical motor regions and the cerebellum previously. Despite this, studies investigating $M1$-cerebellum connectivity have been conducted using resting-state fMRI [6]. Attempts to conclude the efficacy of modulation between cortical and cerebellum are hampered by the limited number of studies that have been conducted [20]. In this research, participants had the potential to upregulate the activity level between M1-cerebellum connectivity since the correlation in the last run was stronger than that of the first.
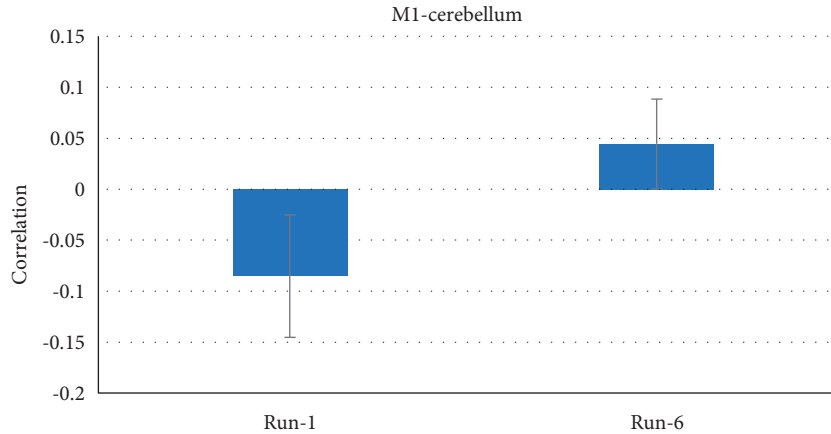
FIGURE 3: This graph shows the connectivity between $M1$ and the cerebellum in the first and final runs. The standard deviation of the mean is represented by the error bars.

TABLE 2: Clusters of brain activations seen during the NF modulation.

Cortex	X	Y	Z	t	P-value	Number of voxels
RT parietal lobe, BA 7	21	−58	52	23.923	0.00174	5502
RT insula, BA 13	32	−22	13	109.41	0.00008	14583
RT caudate body	12	7	25	52.201	0.00037	1470
LT parietal lobe, BA 7	−21	−58	34	54.443	0.000337	1278
LT precentral gyrus, BA 4	−33	−19	55	17.8134	0.003137	257
LT insula, BA 13	−33	8	10	92.853	0.000116	3903
LT frontal lobe, BA 9	−39	47	28	47.463	0.000444	2581
LT cerebellum, anterior lobe	−32	−58	−23	23.6665	0.00178	646
RT cerebellum, posterior lobe	33	−64	−24	9.0189	0.002878	308
RT medial frontal gyrus, BA 6	15	−7	56	24.7062	0.000145	2221
LT superior frontal gyrus, BA 6	−9	−1	63	7.4269	0.00505	603
RT cerebellum, anterior lobe	21	−59	−23	23.9235	0.00174	5502
LT cerebellum, posterior lobe	−12	−59	−14	31.8582	0.00098	520

A linear regression of average connectivity values was utilized to measure the upregulation over neurofeedback runs.

The results indicate that $M1$-cerebellar connectivity seems to be increased by time. However, t-test indicates that the upregulating is not statistically significant. Furthermore, through linear regression, the NF group's mean $M1$-cerebellum activity was not apparent. This means that there was no significant relationship between NF and M1-cerebellum activity. In other words, the findings suggest that being NF does not seem to have an impact on cerebellum activity. This suggests that increased $M1$-cerebellum connectivity was likely a product of neurofeedback modulation and not intrinsic changes made by participants after putting more effort to enhance the modulation.

This study has several implications for clinicians and scientists who are investigating interactions between brain function and mental disorders. First, it suggests that further research is needed in this area before any firm conclusions can be drawn about whether or not being NF affects cerebellar function. Second, it highlights the importance of using appropriate methodology when studying these types of relationships—particularly when looking at $M1$-cerebellum activation levels.

A t-test was used as an exploratory examination to contrast the correlation between the first and final runs. Despite this, neither test resulted in a statistically significant result. This study recruited four participants at a time (four). This study's impact is modest because bigger sample sizes result in more effects. Motor imagery (MI) is a condition in which a person performs mental motor tasks but has no physical motor output [21]. According to several prior studies, MI seems to activate comparable motor areas as those engaged in preparation for and actual execution of actions [22]. Mental tasks in motor imagery have modulated brain activity in specific areas such as $M1$ [23] and SMA [24].

The study by [12, 14] showed that cerebellar activation and M1 activation resulted from motor imagery tasks. These findings are consistent with those of our research because they showed cerebellar activation and $M1$ activation resulting from the same type of task. This is an important finding because it supports the notion that the brain responds similarly to different types of mental activities. The authors note that this could be due to the encoding of movement representations within cerebellum and muscle cortex, respectively. This may allow for improved coordination or speed during movements downstream from these areas. Additionally, the increased cortical activity might
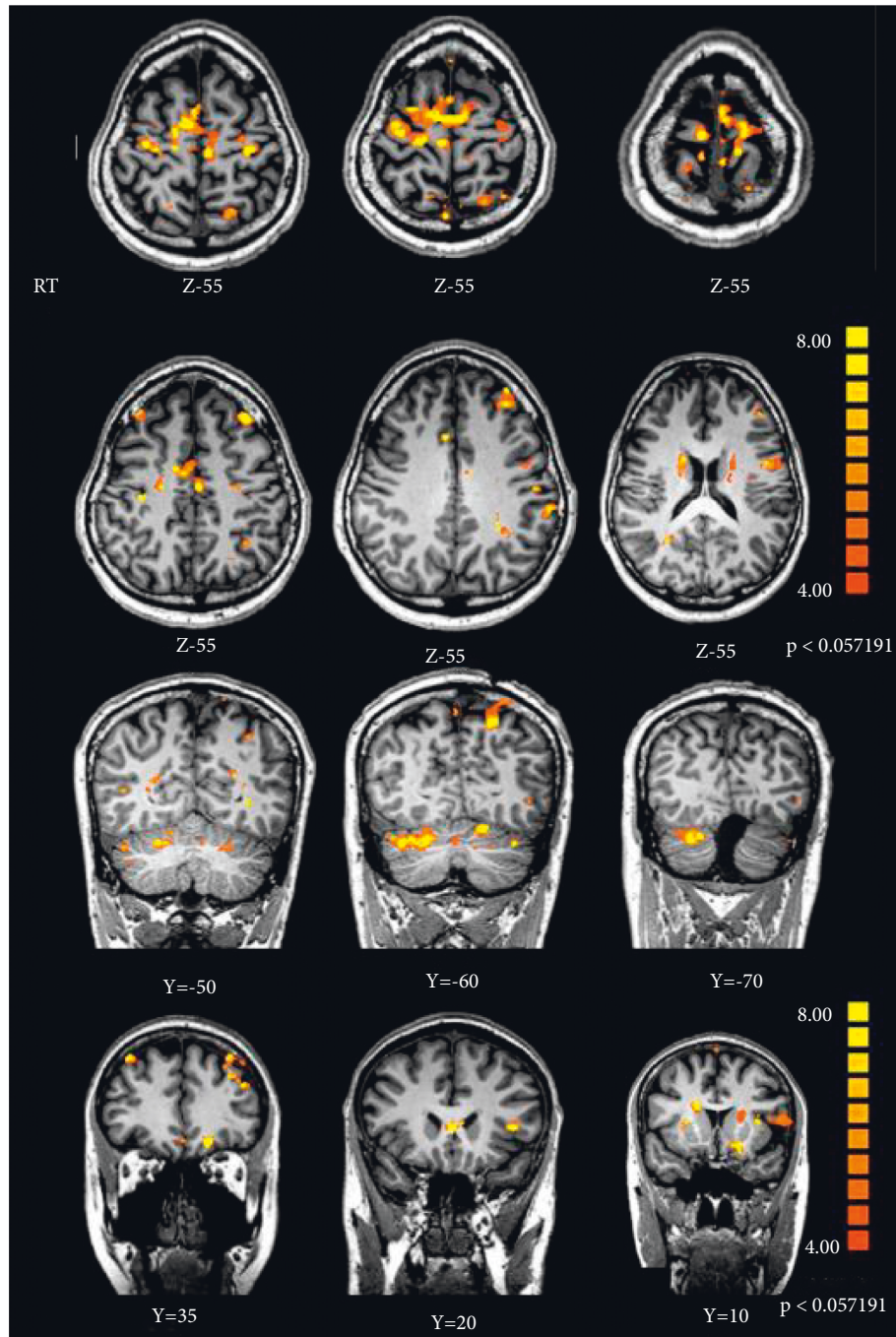
FIGURE 4: For the NF group (coronal images), the results of the NF runs analysis are displayed. At $p < 0.01$ (uncorrected), these activations are significant.

account for the subjective experience of imaging motor actions while performing the experiment.

In this rt-fMRI NF research, the implicit strategy is often employed in which participants were instructed to search for a strategy that enhances their activation score during the experiment [25]. Participants reported that actions such as lifting weights or playing music increased the thermometer bar during the neurofeedback experiment. However, the downside of the implicit strategy is participants may quit exploring new mental actions if they discover one that boosts their neural activity level [25].

Whole-brain activity and brain-distributed learning processes can be monitored and investigated using real-time fMRI. Regarding most real-time fMRI investigations, the training effects are mostly seen in the neurofeedback target region [23]. This may be because this is where the focus of attention typically lies during cognitive tasks, such as those used in neurofeedback training. It has been shown that repeated exposures to a stimuli or task over time can lead to an enhanced response within that particular region of the brain. However, some investigations report increased activation throughout the whole brain during neurofeedback

training, which might be due to general enhancements of cognitive function that result from neurofeedback training or specific neural changes induced by neurofeedback stimulation in certain regions. Future research should explore this question more closely and determine which regions of the brain may benefit from neurofeedback intervention most specifically. The influence of self-regulation and learning of self-regulatory strategies on a successful connection has been investigated [29]. According to numerous real-time fMRI NF research, neurofeedback training affects distinct changes in target region connectivity, predominantly strengthening important connections while weakening others. Consequently, whole-brain analysis permits researchers to see how activation and network alterations occur in specific regions. The full-brain examination was used during the training process to see how the brain became activated.

Bilateral parietal lobe activation was seen during NF training. The processing of visual stimuli could have triggered these areas. Participants got visual feedback from a thermometer bar to show the activation levels of a motor imagery task throughout the experiment. These feedback projections may boost parietal lobe recruitment via real-time neurofeedback [30]. There has been a lot of recent research exploring how different types of imagery (e.g., visual, auditory, olfactory, etc.), and in particular motor imagery, can influence various aspects of cognitive function.

In a recent study, researchers looked at how motor imagery might be related to brain activation. They hypothesized that using motor imagery (MI) might be linked to fMRI activation in the left parietal lobe [31, 32]. The insula may be activated during NF training because it processes intentional actions [33]. Furthermore, it's possible that using motor imagery (MI) is linked to fMRI activation in the left parietal lobe. The insula may be activated during NF training because it processes intentional movement and awareness [36]. In addition, activation in the sensorimotor-related brain regions, such as the insula, may be linked to self-agency perception during voluntary motor activities [32, 36]. The insula is also a critical component of the brain, as Emmert et al. [36] demonstrated during real-time fMRI neurofeedback.

Real-time fMRI NF may alter M1-cerebellum connectivity, according to current findings. The complete-brain analysis's activation in the whole brain and the cerebellum bolstered this conclusion. Lindeman and colleagues were able to modify M1 and the cerebellum using a non-MRI approach. They activated Purkinje neurons using whiskers to alter the coherence between S1 and M1. We recommend doing more research with a bigger sample size in the future because our online and offline findings suggest an impact on targeted connectivity [37].

5. Conclusion

Neurofeedback is a new and advance method that has been found to help people with various mental health conditions, such as anxiety and depression. Neurofeedback is based on the theory that your brain functions in much the same way as a computer—you can train it to respond in a certain way by providing feedback to it. By using neurofeedback, patients can learn how to control their motor, thoughts, and emotions and improve their overall mental health and motor outcome.

The cerebellum plays an important role in skilled movement timing and precision. M1 is critical in addition to planning and carrying out actions. The possibility of changing the M1-cerebellar connection was highlighted using real-time fMRI. We used this method in our research and found that each participant's neurofeedback connection between M1 and the cerebellum appeared to strengthen. Nevertheless, this improvement was not statistically significant. Further investigation, the recruitment of additional participants, and forming a control group are advised.

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] J. A. Bernard, R. D. Seidler, K. M. Hassevoort et al., "Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches," *Frontiers in Neuroanatomy*, vol. 6, p. 31, 2012.

[2] J. P. Kuhtz-Buschbeck, C. Mahnkopf, C. Holzknecht, H. Siebner, S. Ulmer, and O. Jansen, "Effector-independent representations of simple and complex imagined finger movements: a combined fMRI and TMS study," *European Journal of Neuroscience*, vol. 18, no. 12, pp. 3375–3387, 2003.

[3] P. T. Fox, M. E. Raichle, and W. T. Thach, "Functional mapping of the human cerebellum with positron emission tomography," *Proceedings of the National Academy of Sciences of the U S A*, vol. 82, no. 21, pp. 7462–7466, 1985.

[4] C. J. Stoodley and J. D. Schmahmann, "Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies," *NeuroImage*, vol. 44, no. 2, pp. 489–501, 2009.

[5] J. E. Hoover and P. L. Strick, "The organization of cerebellar and basal ganglia outputs to primary motor cortex as revealed by retrograde transneuronal transport of herpes simplex virus type 1," *Journal of Neuroscience*, vol. 19, no. 4, pp. 1446–1463, 1999.

[6] Z. J. Daskalakis, G. O. Paradiso, B. K. Christensen, P. B. Fitzgerald, C. Gunraj, and R. Chen, "Exploring the connectivity between the cerebellum and motor cortex in humans," *The Journal of Physiology*, vol. 557, no. 2, pp. 689–700, 2004.

[7] M. Manto, J. M. Bower, A. B. Conforto et al., "Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement," *The Cerebellum*, vol. 11, no. 2, pp. 457–487, 2012.

[8] S. Arnold Anteraper, X. Guell, A. D'Mello, N. Joshi, S. Whitfield-Gabrieli, and G. Joshi, "Disrupted cerebrocerebellar intrinsic functional connectivity in young adults with high-functioning autism spectrum disorder: a data-driven, whole-brain, high-temporal resolution functional magnetic

resonance imaging study," *Brain Connectivity*, vol. 9, no. 1, pp. 48–59, 2019.

[9] O. Kaut, C. Mielacher, R. Hurlemann, and U. Wüllner, "Resting-state fMRI reveals increased functional connectivity in the cerebellum but decreased functional connectivity of the caudate nucleus in Parkinson's disease," *Neurological Research*, vol. 42, no. 1, pp. 62–67, 2020.

[10] T. C. Ramos, J. B. Balardin, J. R. Sato, and A. Fujita, "Abnormal cortico-cerebellar functional connectivity in autism spectrum disorder," *Frontiers in Systems Neuroscience*, vol. 12, p. 74, 2018.

[11] S. H. Mostofsky, S. K. Powell, D. J. Simmonds, M. C. Goldberg, B. Caffo, and J. J. Pekar, "Decreased connectivity and cerebellar activity in autism during motor task performance," *Brain*, vol. 132, no. 9, pp. 2413–2425, 2009.

[12] M. L. Blefari, J. Sulzer, M. C. Hepp-Reymond, S. Kollias, and R. Gassert, "Improvement in precision grip force control with self-modulation of primary motor cortex during motor imagery," *Frontiers in Behavioral Neuroscience*, vol. 9, p. 18, 2015.

[13] A. Guillot, C. Collet, V. A. Nguyen, F. Malouin, C. Richards, and J. Doyon, "Brain activity during visual versus kinesthetic imagery: an fMRI study," *Human Brain Mapping*, vol. 30, no. 7, pp. 2157–2172, 2009.

[14] M. Lotze and U. Halsband, "Motor imagery," *Journal of Physiology Paris*, vol. 99, no. 4-6, pp. 386–395, 2006.

[15] R. Gentili, C. E. Han, N. Schweighofer, and C. Papaxanthis, "Motor learning without doing: trial-by-trial improvement in motor performance during mental training," *Journal of Neurophysiology*, vol. 104, no. 2, pp. 774–783, 2010.

[16] L. O. Solstrand, O. Lungu, and J. Doyon, "Cerebellar contribution to motor and non-motor functions in Parkinson's disease: a meta-analysis of fMRI findings," *Frontiers in Neurology*, vol. 11, p. 127, 2020.

[17] R. C. Oldfield, "The assessment and analysis of handedness: the Edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, 1971.

[18] N. Callow and R. Roberts, "Imagery research: an investigation of three issues," *Psychology of Sport and Exercise*, vol. 11, no. 4, pp. 325–329, 2010.

[19] A. Isaac, D. F. Marks, and D. G. Russell, "An instrument for assessing imagery of movement: the Vividness of Movement Imagery Questionnaire (VMIQ)," *Journal of Mental Imagery*, vol. 10, no. 4, 1986.

[20] T. Hanakawa, "Rostral premotor cortex as a gateway between motor and cognitive networks," *Neuroscience Research*, vol. 70, no. 2, pp. 144–154, 2011.

[21] S. J. Johnston, S. G. Boehm, D. Healy, R. Goebel, and D. E. Linden, "Neurofeedback: a promising tool for the self-regulation of emotion networks," *NeuroImage*, vol. 49, no. 1, pp. 1066–1072, 2010.

[22] R. Goebel, F. Esposito, and E. Formisano, "Analysis of functional image analysis contest (FIAC) data with brain-voyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis," *Human Brain Mapping*, vol. 27, no. 5, pp. 392–401, 2006.

[23] A. Tursic, J. Eck, M. Lührs, D. E. Linden, and R. Goebel, "A systematic review of fMRI neurofeedback reporting and effects in clinical populations," *NeuroImage: Clinic*, vol. 28, Article ID 102496, 2020.

[24] B. Cengiz and H. E. Boran, "The role of the cerebellum in motor imagery," *Neuroscience Letters*, vol. 617, pp. 156–159, 2016.

[25] S. S. Yoo, J. H. Lee, H. O'Leary, L. P. Panych, and F. A. Jolesz, "Neurofeedback fMRI-mediated learning and consolidation of regional brain activation during motor imagery," *International Journal of Imaging Systems and Technology*, vol. 18, no. 1, pp. 69–78, 2008.

[26] R. deCharms, K. Christoff, G. H. Glover, J. M. Pauly, S. Whitfield, and J. D. Gabrieli, "Learned regulation of spatially localized brain activation using real-time fMRI," *NeuroImage*, vol. 21, no. 1, pp. 436–443, 2004.

[27] S. Al-Wasity, S. Vogt, A. Vuckovic, and F. E. Pollick, "Upregulation of supplementary motor area activation with fMRI neurofeedback during motor imagery," *eNeuro*, vol. 8, no. 1, p. ENEURO.0377, 2021.

[28] C. Paret, N. Goldway, C. Zich et al., "Current progress in real-time functional magnetic resonance-based neurofeedback: methodological challenges and achievements," *NeuroImage*, vol. 202, Article ID 116107, 2019.

[29] G. Rota, G. Handjaras, R. Sitaram, N. Birbaumer, and G. Dogil, "Reorganization of functional and effective connectivity during real-time fMRI-BCI modulation of prosody processing," *Brain and Language*, vol. 117, no. 3, pp. 123–132, 2011.

[30] S. Lindeman, S. Hong, L. Kros et al., "Cerebellar Purkinje cells can differentially modulate coherence between sensory and motor cortex depending on region and behavior," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 2, Article ID e2015292118, 2021.

[31] P. Andersson, F. Ragni, and A. Lingnau, "Visual imagery during real-time fMRI neurofeedback from occipital and superior parietal cortex," *NeuroImage*, vol. 200, pp. 332–343, 2019.

[32] F. Lebon, U. Horn, M. Domin, and M. Lotze, "Motor imagery training: kinesthetic imagery strategy and inferior parietal f MRI activation," *Human Brain Mapping*, vol. 39, no. 4, pp. 1805–1813, 2018.

[33] S. Tinaz, K. Para, A. Vives-Rodriguez et al., "Insula as the interface between body awareness and movement: a neurofeedback-guided kinesthetic motor imagery study in Parkinson's disease," *Frontiers in Human Neuroscience*, vol. 12, p. 496, 2018.

[34] C. Farrer and C. D. Frith, "Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency," *NeuroImage*, vol. 15, no. 3, pp. 596–603, 2002.

[35] B. Lorey, M. Bischoff, S. Pilgramm, R. Stark, J. Munzert, and K. Zentgraf, "The embodied nature of motor imagery: the influence of posture and perspective," *Experimental Brain Research*, vol. 194, no. 2, pp. 233–243, 2009.

[36] K. Emmert, "Improvement of Real-Time fMRI Neurofeedback for Clinical Applications," 2016.

[37] Y. Madkhali, "Real-time fMRI Connectivity Neurofeedback for Modulation of the Motor System," Doctoral dissertation, University of Glasgow, Glasgow, Scotland, 2022.

Research Article

Breast Cancer Pathological Image Classification Based on the Multiscale CNN Squeeze Model

Yahya Alqahtani [1], **Umakant Mandawkar** [2], **Aditi Sharma** [3], **Mohammad Najmus Saquib Hasan** [4], **Mrunalini Harish Kulkarni** [5], and **R. Sugumar** [6]

[1]*Faculty of Computer Science and Information Technology, Jazan University, Jizan, Saudi Arabia*
[2]*SVKM'S Institute of Technology, Dhule, India*
[3]*Department of Computer Science Engineering & Information Technology,*
Institute of Engineering & Technology (An Autonomous Constituent Institute of Dr. A.P.J. Abdul Kalam Technical University),
UP, Lucknow, India
[4]*Wollega University, Nek'emtē, Ethiopia*
[5]*Department of Pharmacy, School of Pharmacy, Vishwakarma University, Pune, India*
[6]*Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical*
and Technical Sciences, Chennai 602105, India

Correspondence should be addressed to Mohammad Najmus Saquib Hasan; mohammadk@wollegauniversity.edu.et

The use of an automatic histopathological image identification system is essential for expediting diagnoses and lowering mistake rates. Although it is of enormous clinical importance, computerized breast cancer multiclassification using histological pictures has rarely been investigated. A deep learning-based classification strategy is suggested to solve the challenge of automated categorization of breast cancer pathology pictures. The attention model that acts on the feature channel is the channel refinement model. The learned channel weight may be used to reduce superfluous features when implementing the feature channel. To increase classification accuracy, calibration is necessary. To increase the accuracy of channel recalibration findings, a multiscale channel recalibration model is provided, and the msSE-ResNet convolutional neural network is built. The multiscale properties flow through the network's highest pooling layer. The channel weights obtained at different scales are delivered into line fusion and used as input to the next channel recalibration model, which may improve the results of channel recalibration. The experimental findings reveal that the spatial recalibration model fares poorly on the job of classifying breast cancer pathology pictures when applied to the semantic segmentation of brain MRI images. The public BreakHis dataset is used to conduct the experiment. The network performs benign/malignant breast pathology picture classification collected at various magnifications with a classification accuracy of 88.87 percent, according to experimental data. The diseased images are also more resilient. Experiments on pathological pictures at various magnifications show that msSE-ResNet34 is capable of performing well when used to classify pathological images at various magnifications.

1. Introduction

Breast cancer is the most common cancer in women [1], and the incidence tends to be younger. Pathological detection is regarded as the "gold standard" in the diagnosis of breast cancer [2], and pathological detection is determined by pathology. It is carried out under the microscope, and the

pathological grade is given by the observation of the pathological section. Due to the large variability in the pathological images [3], the observer's experience and subjective differences may affect the most the final diagnosis. Benign tumors are noncancerous. They will not grow or invade surrounding tissue. However, when they form near vital organs, irritate a nerve, or restrict blood flow, they can

be highly dangerous. The majority of benign tumors respond well to treatment. The only method to tell for sure is to get the lump biopsied, even though the fact that tests such as mammograms, ultrasounds, and MRI might provide hints as to whether a mass is malignant.

Malignant tumors are carcinogenic tumors. Our bodies constantly produce new cells to replace worn-out ones. Occasionally, DNA gets damaged during the process, resulting in abnormal cell formation. Instead of vanishing, they continue to expand at a rate that the immunity is unable to keep up with, leading to a tumor. Cancer cells can travel from the tumor to other parts of the body via the circulatory or vascular system. The location of the underlying tumor and whether it has spread are only two of the numerous variables that affect how malignant tumors are treated. Detailed information about the tumor can be revealed by a pathology report to aid with treatment planning, which may involve surgery, radiotherapy, chemotherapy, targeted therapy, and immunotherapy, commonly known as biological therapy.

Automatic classification algorithms for breast cancer pathological images can help pathologists make more accurate diagnoses. The research on breast cancer pathological image classification has made great progress in recent years, and the research methods for this task can be divided into There are 2 categories: one is algorithms based on manual feature descriptors and machine learning, and the other is algorithms based on deep learning.

Utilizing feature descriptors including binary patterns (lbp, gray-scale co-occurrence matrices, and classification techniques) such as random forests and support vector machines, the BreaKHis breast cancer pathology picture dataset [4] was produced. Using a majority vote method, literature [5] integrated the outcomes of each classification algorithm to get at an 87 percent detection accuracy for the same dataset. However, high-quality qualities require specialist knowledge and effort, limiting this technology's application.

Deep learning classification beats standard machine learning classification by adopting a network topology with the convolution layer as the core.

For the first time, literature [6] employs 11-layer and 13-layer deep neural networks to indicate the existence of mitosis in breast cancer histopathology pictures. In total, 14-layer convolutional neural networks have been utilized in the literature [7] to categorize breast pathology images as regular tissue, benign lesions, in situ carcinomas, or invasive malignancy. Literature [8] compared BreaKHis dataset studies utilizing AlexNet-based models. Machine learning categorization enhanced performance by 4–6%. Literature [9] used a magnification-independent deep network to acquire an 83% identification rate; Scheer vector and VGGNet's classification model provide an 87% recognition rate. The work requires preserving enough sturdiness for pathological photos at varied enhancements due to the significant differences in infected pictures at various magnifications. Deep network training needs many training examples, yet pathological breast cancer pictures are scarce. With the growing availability and incorporation of many data types, such as genomics, microarray, and histopathologic data, cancer therapy is moving toward precision medicine. It takes a lot of time and experience to use and analyze a variety of high-dimensional data formats for clinical or translational research jobs. Additionally, combining different data kinds requires more computing power than interpreting each type separately and calls for modeling algorithms that can absorb enormous amounts of complex characteristics. Machine learning algorithms are increasingly being used to automate these processes to help diagnose and detect cancer. Excitingly, DL models may be able to make use of this complexity to present insightful information and find pertinent granular characteristics from a variety of data formats.

To improve the classification model's effectiveness, maximize the few available samples. The channel recalibration model [10] focuses on feature channels. It suppresses superfluous features through learned channel weights and improves classification model performance. A multichannel CNN model was constructed and proposed as a solution to the identified problem using the sensitive lymph node pathologic imaging datasets for breast cancer. The model employs stacked multichannel convolutional units, Internet of Things-based CNN modules, skip cross-layer interconnections, a combination of classical and depth-wise separable convolution layers, and summation and concatenation operations. According to the results, the model does a good job of identifying micrometastases as well as lymph node metastasis.

This article is in the channel to enhance CNN feature utilization. A multiscale channel squeeze-and-excitation (msSE) model is developed according to the refinement model. It uses different max-pooling layers to gather multiscale features; channel recalibration is undertaken on each scale feature. The fused channel weights achieve multiscale channel recalibration for input characteristics. Multiscale features may improve the network's feature information, and channel recalibration can raise the classification model's performance. The network's training set includes breast cancer pathology images at four magnifications, guaranteeing the classification model is robust to multiple embellishments and meets clinical expectations.

2. Related Work

2.1. Residual Structure. One illustration of a multilayer neural network's unstable behavior is the vanishing gradient problem (VGP). Networks are unable to return gradient information to the model's input layers. Gradients for deeper layers in a multilayer network are calculated as the sum of many gradients of activation functions. These gradients will quickly disappear when they are tiny or zero. On the other hand, if they are more than one, it may explode. As a result, updating and computing become quite difficult. The partial derivatives for the variables of the NN, which are the gradient's constituent elements, grow exponentially tiny in the VGP, practically negligibly changing the variables with the gradient.

The vanishing gradient problem of deep convolution neural networks makes it difficult to train deeper network models. The residual structure proposed literature [11]

solves this problem and enables deeper convolutional neural networks to be trained efficiently. The structure of regression is shown in Figure 1.

The calculation process of the residual structure can be expressed as

$$y = F(x) + x, \tag{1}$$

where x is the input feature of the convolutional layer, y is the output of the residual structure feature, and $F(x)$ is the result after the convolutional layer mapping.

Suppose the residual structure is expected to fit the mapping as $H(x)$. Due to the existence of the additional equivalent mapping, the mapping to be fitted by the convolutional layer in the residual structure becomes the mapping with residual $(x) = H(x) - x$. This is easier to learn than the original expected fitting mapping. The residual structure does not introduce additional parameters and can be trained through backpropagation. The residual network with the residual structure as the main body increases the number of network layers at the same time, which can avoid the gradient vanishing problem.

2.2. Channel Recalibration Model. The attention model [12, 13] was first applied to natural language processing, by introducing attention weights to make the network model "attention" to useful In recent years, attention models have been applied to the field of computer vision [14, 15], by suppressing the uninteresting regions in the feature map, the network's attention is focused on the region of interest. Different from focusing on the feature map The channel refinement model is a model of attention acting on the channel domain of the feature map. It is proposed in SENet [10] designed by the ILSVRC17 competition classification task champion. The channel recalibration model weights the input features by channel, so that the network's attention is focused on useful features, and the channel weights can be learned through training. The channel recalibration model can be combined with VGGNet [16], ResNet [11], GoogLeNet [17], and other networks, and the residual structure and join The SE residual structure of channel recalibration is shown in Figure 2.

The channel recalibration model squeezes the input feature U in channel order according to:

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j), \tag{2}$$

where z_c is the result of squeezing the feature of the first channel in the input feature; $F_{sq}(*)$ is the squeeze function; u_c is the feature of the first channel c in the input feature, and H and W are the height and width, respectively; $u_c(i, j)$ is the value of the feature at u_c the spatial position (i, j). This process can be regarded as a channel-by-channel global pooling operation on the input features.

After the extrusion of each channel feature in the input feature is completed, the weight of each channel is obtained by exciting the extrusion result by the following formula:
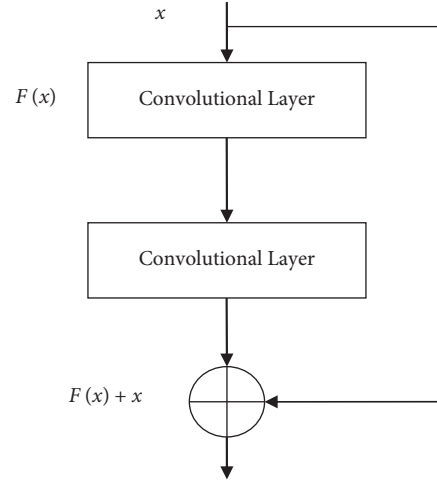
FIGURE 1: Residual structure.

$$s = F_{ex}(z, W) = \sigma(W_2 \delta(W_1 z)), \tag{3}$$

where s is the weight of the feature channel; $F_{ex}(*, *)$ is the excitation function; z is the result of extruding the feature; $\sigma(*)$ is the sigmoid function; $\delta(*)$ is the ReLU function [18]; W_1 and W_2 are the weights of the two fully connected layers FC, respectively.

The first fully connected layer in the excitation process converts the number of feature channels depending on c reduced to c/r, where the compression ratio, and the output is only retaining values greater than zero after the ReLU function. The second fully connected layer restores the number of feature channels to c, so as to be consistent with the number of channels of the input feature. The final weight is obtained through the sigmoid function and the limit is 0~1.0:

$$\ddot{x}_c = F_{\text{scale}}(u_c, s_c) = s_c u_c. \tag{4}$$

In the above formula: $\ddot{x}_c$ is the output characteristic of the channel after the recalibration of the channel feature; is the weight of the cth channel in the input feature; $F_{\text{scale}}(*, *)$ is a scaling function, which is used to multiply the features of a specific channel with the corresponding channel weight. Equation (4) realizes the recalibration of the feature channel by multiplying the feature of a specific channel with the corresponding channel weight, and the whole process suppresses the features that are useless to the classification result, thereby improving the classification accuracy.

3. Proposed Algorithm Description

On the basis of the channel recalibration model, the input of the channel recalibration model is changed from single-scale features to multiscale features, and the feature channel weights learned at each scale are fused to obtain the final feature channel weights. Add multiscale The msSE residual structure of the channel recalibration model is shown in Figure 3. Convolutional neural networks using multiscale features are often used in tasks such as target detection and
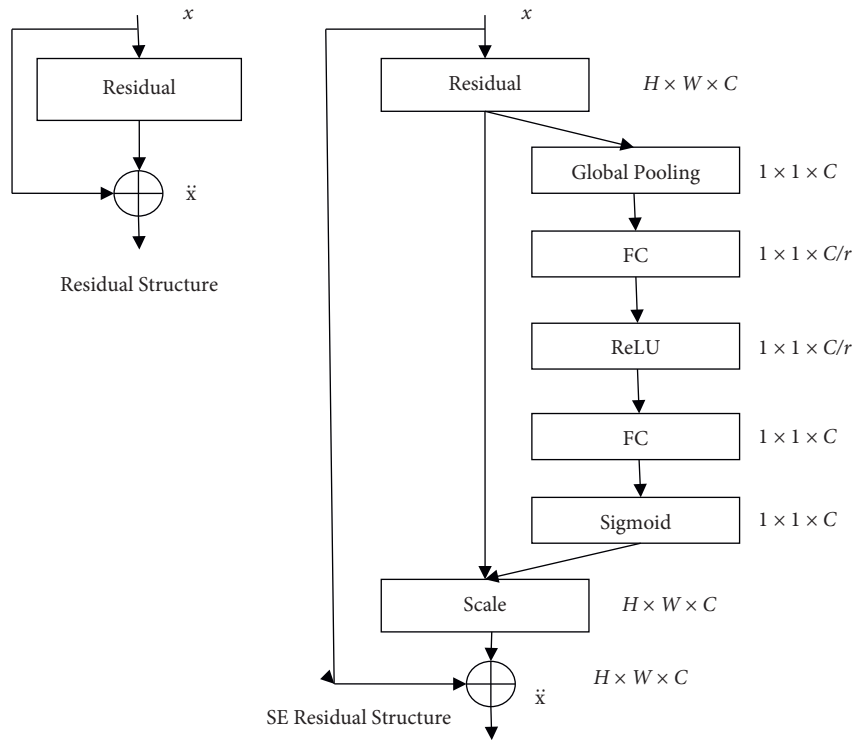
Figure 2: Residual structure and SE residual structure.
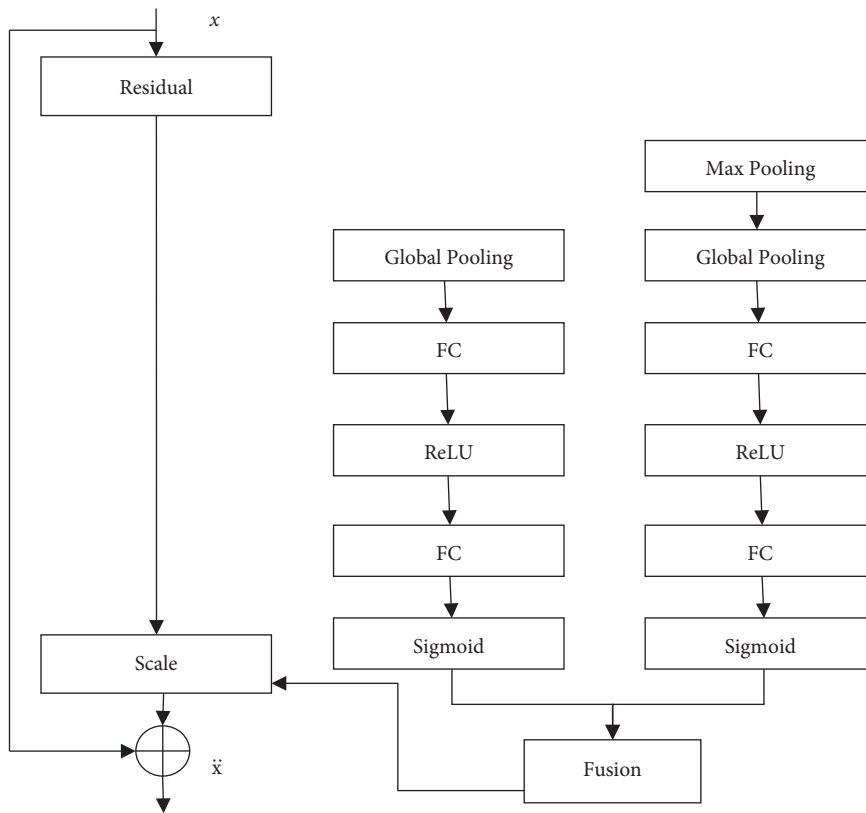
Figure 3: msSE residual structure.

recognition [15, 19–21] and image semantic segmentation [22–24]. Using feature information at multiple scales can make the final result more accurate. In Figure 3, the multiscale features are combined with spatial pooling gold.

Pyramid [19] similar structure is obtained: the input features are sent to the max pooling layer with a pooling kernel size of 2×2 and a pooling stride of 2 to obtain features of another scale. To obtain more scales, features can be achieved by changing the number of max pooling layers and related parameters. The reason for using max pooling layers to obtain multiscale features is that the max pooling operation can retain the most significant feature information and corresponding spatial information in the feature map; the maximum. The pooling layer has no model parameters that need to be learned and can achieve multiscale features while ensuring that as little computation as possible is introduced into the network. In Figure 3, fusion represents the fusion process of channel weights. Although channel attention appears to be economical in terms of variables and FLOPs overhead, one key problem is the scaling process, which involves broadcasting the weighted channel vector and applying by multiplying it element-wise to the input tensor. This intermediary broadcasted tensor occupies the same dimension as the input, resulting in a significant rise in memory complexity. As a result, the training process becomes slower and more memory intensive. The approach is highly expensive and augments the original model with a sizable number of parameters and FLOPS. Although in the big scheme of things, this overhead could be relatively little, there have been numerous novel ways that have outperformed SENets in terms of giving network attention at a very low cost. The method of maximum value and splicing fuses the feature channel weights obtained at different scales.

3.1. Additive Fusion.
In Figure 3, the channel weight obtained by the additive fusion method 2 is the element-by-element addition of the channel weights under the two feature scales, and then the obtained weights are multiplied by the input features in the order of the corresponding channels to achieve multiscale channel recalibration. A process is as follows:

$$\dot{U}_{2\text{way\_add}} = \left(S_{c0} + S_{c1} \right) U_{s0}. \tag{5}$$

In the formula $\dot{U}_{2\text{way\_add}}$ for 2 the results of multiscale channel recalibration using additive fusion at each feature scale, U_{s0} is the input feature, and S_{c0} is the channel weight of the input feature, S_{c1} is the channel weight at another scale.

3.2. Maximum Fusion.
Unlike additive fusion, maximum fusion selects specific channels 2 the maximum value of the weight under each scale is used as the weight of the channel. At this time, the multiscale channel recalibration process is as follows:

$$\dot{U}_{2\text{way\_max}} = \left(S_{c0}, S_{c1} \right) U_{s0}. \tag{6}$$

In the formula $\dot{U}_{\_(2\text{way\_max})}$ for 2, the results of multiscale channel recalibration using maximum fusion at each feature scale; $\max(*, *)$ is the maximum function.

Select 2, respectively, in channel order the maximum value of the channel weight under each scale is used as the weight of the channel.

3.3. Splicing Fusion.
When there are two scale features, the splicing fusion method first splices the channel weights at each scale according to a specific coordinate axis, and then maps the result to the final channel weight through the subsequent convolution layer. The channel weight size is $N \times C \times 1 \times 1$, where the batch image size is and the number of channels of the input feature is C The specific implementation of splicing fusion can be divided into the following two types according to the selection of the splicing coordinate axis.

(a) Take the second coordinate axis (axis1) as the splicing coordinate axis, denoted as cat1. At this time, the multiscale channel recalibration process can be expressed as

$$\dot{U}_{2\text{way\_cat1}} = F_{\text{conv1}} \left(S_{c\_\text{cat1}} \right) U_{s0}, \tag{7}$$

where $\dot{U}_{2\text{way\_cat1}}$ is the result of multiscale channel recalibration achieved by splicing and fusion cat1 at 2 scales, S_{c_cat1} is the result of splicing the channel weights obtained according to the second coordinate axis, size is $N \times 2C \times 1 \times 1$, F_conv1 (*) for the convolutional layer conv1 the mapping function, the size of the convolution kernel is 1×1, the number of input channels is 2C, and the number of output channels is C.

(b) Take the third coordinate axis (axis2) as the splicing coordinate axis, denoted as cat2. At this time, the multiscale channel recalibration process can be expressed as

$$\dot{U}_{2\text{way\_cat2}} = F_{\text{conv2}} \left(S_{c\_\text{cat2}} \right) U_{s0}, \tag{8}$$

where $\dot{U}_{2\text{way\_cat2}}$ is the result of multiscale channel recalibration realized by splicing and fusion cat2 at two scales, S_{c_cat2} is the result of splicing the channel weights obtained in the two scales according to the third coordinate axis, size is $N \times 2C \times 2 \times 1$, F_{conv2} (*) is the mapping function of the convolution layer, where the size of the convolution kernel is 2×1, The number of input and output channels are both C.

4. Experimental Results and Analysis

4.1. Dataset.
Experimental dataset: BreaKHis dataset, includes 7909 breast cancer pathology pictures from 82 individuals (24 benign and 58 malignant). The 700460 pathological photos in the dataset have 4 magnifications (40x, 100x, 200x, 400x). Table 1 shows the dataset's picture distribution. Figure 4 shows benign/malignant breast tumors from BreaKHis.

(a)   (b)
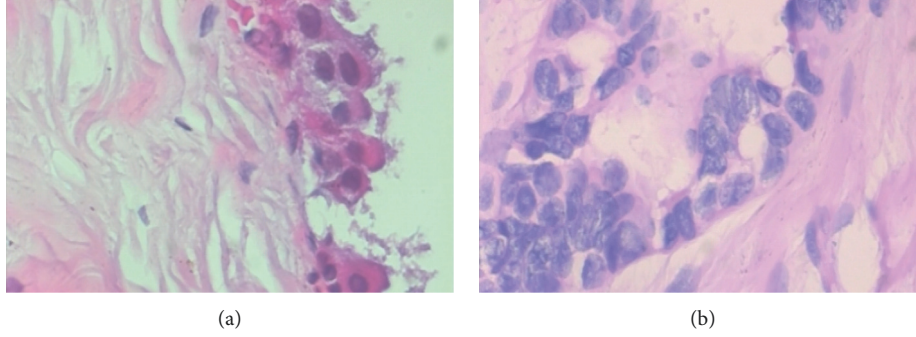
FIGURE 4: Benign and malignant breast tumor images. (a) Benign breast tumor image. (b) Malignant breast tumor image.

TABLE 1: Distribution of pictures under different magnifications and categories.

Gain	Number of tumor images		
	Benign	Malignant	Total
40 Times	750	1644	2394
100 Times	773	1725	2498
200 Times	748	1668	2416
400 Times	706	1479	2184

4.2. Experimental Environment and Settings. The accuracy Acc (accuracy), the precision rate Pr (precision), the recall rate *R* (recall), and the area under the ROC curve AUC are used as the measurement indicators of the classification results.

The formulas for calculating the rate and recall rate are as follows:

$$A_{cc} = \frac{TP + TN}{TP + FP + TN + FN'},$$

$$P_r = \frac{TP}{TP + FP'}, \tag{9}$$

$$R = \frac{TP}{TP + FN}.$$

In the above equation, TP represents the true positive example, FP represents the false positive example, and TN represents the true negative example, and FN means false negative example.

The dataset was not augmented in the experiment. The ratio of the training set to the test set was divided into 85% and 15%, and the images included were randomly selected at the beginning of the training. All comparative experiments used official source code or public code, all the network model uses the same image preprocessing method and training settings, and the experimental data of each network is obtained by averaging the results obtained by 5 times of training.

The training images are preprocessed as follows: (1) The image size is adjusted to a fixed 224×224; (2) The image is randomly rotated by 90°; (3) The brightness, contrast, saturation, and chroma of the image are randomly fine-tuned, which makes the training. The network can be more robust to the staining differences between pathological images; (4) normalize the images. Different from the training set, the

preprocessing method of the test set images only includes adjusting the image size to 224×224 with normalization chemical processing.

The initial parameters of all network models in the experiment are obtained by random initialization, and the loss function is binary cross entropy, using momentum. The weights are updated using the network's stochastic gradient (SGD) algorithm. The momentum value is set to 0.9 and the starting training rate is set at 0.0001. The network's retraining batch image size is 64, whereas the test batch picture size is 128. Res-Net18-based network the number of training iterations is 10,600, and the number of network training iterations based on Res-Net34 is 21,200. If the test accuracy does not improve after every 1060 iterations, the learning rate is reduced to 0.1 times before.

Since the multiscale features are obtained by down-sampling the input features and the convolutional layer in the network will reduce the feature size, in order to keep the feature size in a reasonable range, the experiment only selects networks with feature scales of 2 and 3 for experiments. Select the network with 2 feature scales as msSE-ResNet-2way, and the network with 3 feature scales as msSE-ResNet-3way. In splicing fusion, the network that applies the sigmoid function to the output is added after its name adds sigma to differentiate.

4.3. Experiments Based on ResNet18. Comparison and analysis of msSE-ResNet18 and other networks, the experimental results of each network on the test set are shown in Table 2, and the ROC curve is shown in Figure 5. In the figure, FPR is the ratio of false positives, and TPR is true positives the classification results of breast cancer pathological images of different magnifications in the test set by each network in the experiment are shown in Table 3. Table 4 shows the comparison of magnification-related classification results for all networks.

It can be seen from Table 2 that the test accuracy of ResNet18 is 84.53%, which is higher than the 83.56% of SE-ResNet18. Literature [16] proposed a spatial channel recalibration model (spatial and channel Squeeze-and-Excitation, scSE), which simultaneously performs spatial and channel recalibration, and use the maximum value of the weights obtained by the two as the feature channel weight.
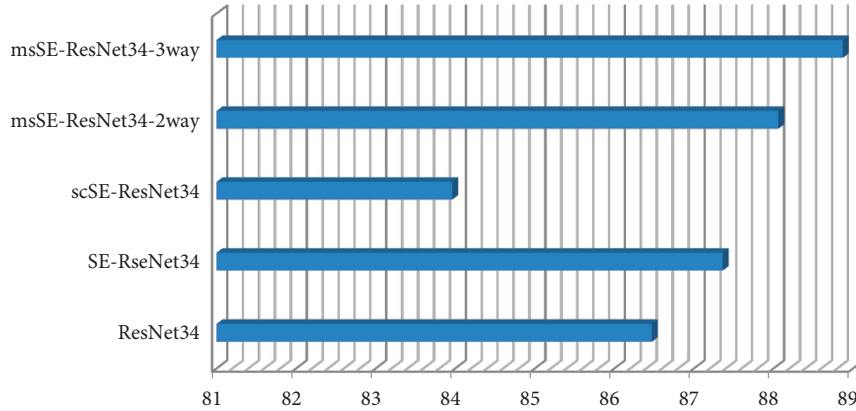
FIGURE 5: Comparison of accuracy between msSE-ResNet34 and other networks.

TABLE 2: Comparison of msSE-ResNet18 and other networks' categorization outcomes.

Model	A_{cc}/%	AUC
ResNet18	84.53	0.8878
SE-ResNet18	83.56	0.8791
scSE-ResNet18	83.90	0.8677
msSE-ResNet18-2way	86.81	0.9266
msSE-ResNet18-3way	86.00	0.9107

TABLE 3: Comparison of magnification-related categorization results for all networks.

Model	Number
ResNet18	1
SE-ResNet18	2
scSE-ResNet18	3
msSE-ResNet18-2way	4
msSE-ResNet18-3way	5

The accuracy of scSE-ResNet18 is 83.90%. The accuracy of msSE-ResNet18-2way reaches 86.81%, and the test of msSE-ResNet18-3way the accuracy is 86%, which is significantly improved compared with other networks. The comparison of magnification-related classification results for all networks is shown in Figure 4.

The ROC curve in Figure 5 further reflects the classification performance of each network. The AUC of msSE-ResNet18 with two scales is above 0.9, achieving better performance than other networks.

The experimental findings are shown in Table 3, and they show that msSE-ResNet18-2way retains strong resilience to diseased pictures at various magnifications; at magnifications greater than 40 times, msSE-Res-Net18-3way has the same performance as msSE-ResNet18-2way comparable classification performance. Since the task of the experiment is to classify benign/malignant breast pathological images, the classification accuracy is more important. The msSE-ResNet with multiscale channel recalibration is obtained under each magnification. The accuracy rate is higher than that of other comparison networks, which means that msSE-ResNet18 can more accurately find the malignant samples in

the test set, achieve a high recall rate under the premise of ensuring high accuracy, and can find as many positive samples as possible.

The above experimental results show that the multiscale channel recalibration model can recalibrate the input features more accurately by combining the feature information at multiple scales and can improve the performance of the classification model while maintaining robustness to pathological images under different magnifications awesomeness.

Comparison and analysis of msSEResNet18 using different number of feature scales and fusion methods. The experimental results of msSE-ResNet18 using different number of feature scales and fusion methods are shown in Table 3.

Shown in the table, A_{tr} is the training accuracy and A_{te} is the test accuracy.

It can be seen from Table 5 that the highest test accuracy under the two scales is relatively close, the channel weights under the two scales are suitable for fusion by the linear addition method; the channel weights under the three scales are suitable for selecting the splicing fusion method. Splicing the selection of the splicing coordinate axis in the fusion will affect the accuracy by about 1%. At this time, splicing with the second coordinate axis (cat1) can achieve higher accuracy. In addition, applying the sigmoid function to the output channel weight of the splicing fusion will significantly degrade the classification performance, because the sigmoid function greatly limits the range of values for the channel weights learned by the convolutional layers.

4.4. Experiments Based on ResNet34

4.4.1. Comparison and Analysis of msSE-ResNet34 and Other Networks. The experimental results of each network on the test set in the experiment are shown in Table 5, and the ROC curve is shown in Figure 6. As shown in Table 6, all the networks in the experiment are different from the test set magnification of pathological images the multiscale channel recalibration model can make the relationship between channels more accurately captured. Experiments on pathological images of different magnifications prove that msSE-esNet34 can be effectively applied to pathological images of different magnifications classification tasks. In Section 4.4.2,

TABLE 4: Comparison of magnification-related classification results for all networks.

Model	40 Times			100 Times			200 Times			400 Times		
	A_{cc}	P_r	R	A_{cc}	P_r	R	A_{cc}	P_r	R	A_{cc}	P_r	R
1	0.822	0.845	0.907	0.836	0.836	0.921	0.864	0.868	0.947	0.875	0.864	0.967
2	0.826	0.820	0.956	0.862	0.861	0.953	0.867	0.862	0.962	0.879	0.865	0.973
3	0.805	0.808	0.941	0.836	0.845	0.935	0.870	0.866	0.962	0.824	0.837	0.918
4	0.862	0.890	0.912	0.862	0.884	0.921	0.880	0.887	0.947	0.889	0.889	0.957
5	0.829	0.856	0.902	0.868	0.878	0.940	0.874	0.905	0.913	0.882	0.884	0.951

TABLE 5: Comparison of categorization results of fusion methods under different number of feature scales.

Number of scales	Fusion method	$A_{tr}\%$	$A_{te}\%$
2	Add	85.42	85.07
2	Max	83.55	82.13
2	Cat1(sign)	84.10	82.96
2	Cat1	85.38	84.64
2	Cat2(sign)	83.93	82.76
2	Cat2	84.61	82.88
3	Add	85.00	83.71
3	Max	83.45	82.22
3	Cat1(sign)	83.76	82.09
3	Cat1	85.61	84.28
3	Cat2(sign)	83.57	82.39
3	Cat2	84.50	83.25
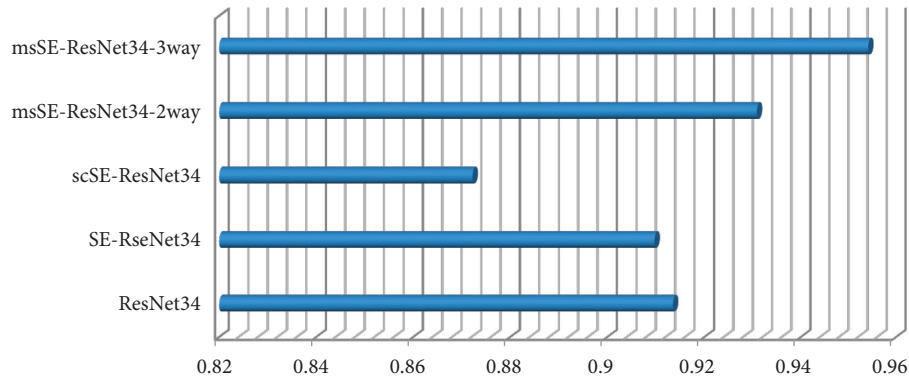
FIGURE 6: Comparison of AUC between msSE-ResNet34 and other networks.

comparison and analysis of msSEResNet34 with different number of feature scales and fusion methods the experimental results of msSE-ResNet34 with different number of feature scales and fusion methods are shown in Table 7.

From Table 7, it can be observed that the highest test accuracy of msSE-ResNet under two different scales is only 0.81% different, with two classification results.

As can be seen from Table 5, with the deepening of the number of ResNet layers, the test accuracy of most networks in the experiment has been greatly improved. SEResNet34 achieved a test accuracy of 87.36%, which is higher than 86.47% of Res-Net34. Applied to the semantic segmentation task of brain MRI images, the experimental results show that the spatial recalibration model does not perform well on the

TABLE 6: Comparison of categorization results between msSE-ResNet34 and other networks.

Model	$A_{cc}\%$	AUC
ResNet34	86.47	0.9135
SE-RseNet34	87.36	0.9097
scSE-ResNet34	83.96	0.8722
msSE-ResNet34-2way	88.06	0.9308
msSE-ResNet34-3way	88.87	0.9541

breast cancer pathological image classification task. The test accuracy of msSE-ResNet34-3way rises to the highest 88.87%, 2 the test accuracy of the network at 1 scale is improved to 88.06%.

TABLE 7: Comparison of magnification-related classification results for all networks.

Model	40 Times			100 Times			200 Times			400 Times		
	A_{cc}	P_r	R	A_{cc}	P_r	R	A_{cc}	P_r	R	A_{cc}	P_r	R
1	0.822	0.845	0.907	0.836	0.836	0.921	0.864	0.868	0.947	0.875	0.864	0.967
2	0.826	0.820	0.956	0.862	0.861	0.953	0.867	0.862	0.962	0.879	0.865	0.973
3	0.805	0.808	0.941	0.836	0.845	0.935	0.870	0.866	0.962	0.824	0.837	0.918
4	0.862	0.890	0.912	0.862	0.884	0.921	0.880	0.887	0.947	0.889	0.889	0.957
5	0.829	0.856	0.902	0.868	0.878	0.940	0.874	0.905	0.913	0.882	0.884	0.951

TABLE 8: Comparison of classification results of each fusion method under different number of feature scales.

Number of scales	Fusion method	$A_{tr}\%$	$A_{te}\%$
2	Add	86.28	86.30
2	Max	85.65	85.97
2	Cat1(sign)	84.88	85.01
2	Cat1	86.86	86.28
2	Cat2(sign)	85.29	85.26
2	Cat2	87.40	85.90
3	Add	85.89	85.43
3	Max	86.44	86.87
3	Cat1(sign)	85.89	85.77
3	Cat1	86.59	86.36
3	Cat2(sign)	85.69	86.54
3	Cat2	87.29	87.09
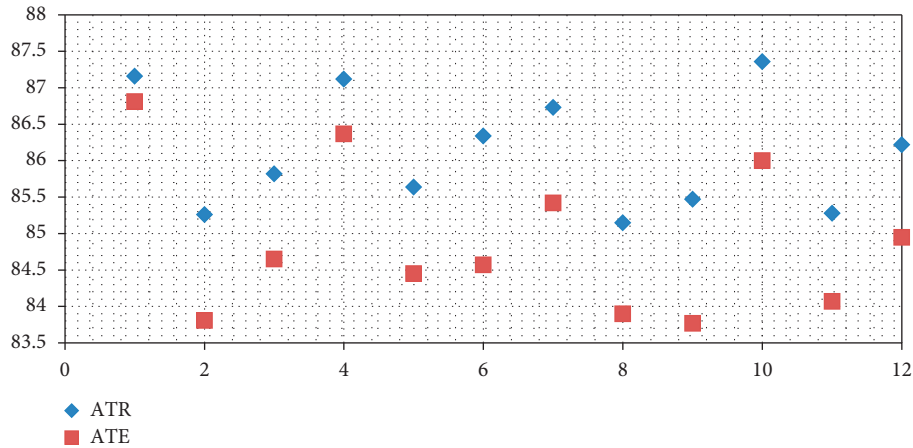
FIGURE 7: Comparison of classification results of each fusion method under different number of feature scales.

Table 6 shows the classification results of all networks related to magnification. It can be seen that the accuracy and accuracy of msSE-ResNet34-3way have been greatly improved at all magnifications, especially at 40 times and can reach a maximum of 90.1% at 400 times. The classification accuracy of msSE-ResNet34-2way has been steadily improved at various magnifications. In the experiments under all magnifications, msSEResNet34 is superior to other comparison networks in both precision and accuracy.

The aforementioned experiments demonstrate that msSEResNet34 can best utilize the rich feature details in the deeper network and that the multiscale channel refinement model can enable a more accurate correlation between channels with the intensifying of the network layers. Experiments on pathological images with multiple magnifications demonstrate that msSE-ResNet34 can be effectively applied to the classification task of pathological images with different magnifications.

4.4.2. Comparison and Analysis of msSEResNet34. msSEResNet34 with the different number of feature scales and fusion methods the experimental results of msSE-ResNet34 with different number of feature scales and fusion methods are shown in Table 7.

It can be seen from Table 8 that the highest test accuracy of msSE-ResNet under two different scales is only 0.81% different, and the performance of additive fusion under two scales is better than that under three scales.

The results obtained in Figure 7, when the network using maximum fusion and splicing fusion at 3 scales is better than 2 scales in classification accuracy. Similar to the conclusion obtained in the experiment based on ResNet18, when there are two scales of features, the addition and fusion of the performance of the method is better than other nonlinear fusion methods, and the splicing fusion method or the maximum fusion method should be preferentially selected under the three scales. The results differ by only about 0.5%, and high classification accuracy can be achieved.

5. Conclusion

The classification job of breast cancer pathology pictures is investigated in this research, and a multiscale channel recalibration model called msSE is proposed, as well as a convolutional neural network called msSE-Res-Net using Res-Net as the network architecture. The fusion of feature weights learned at different scales may significantly increase the dependability of the feature channel weight learning process. Multiscale features can enrich the features in the network information and improve feature usage. The BreaKHis dataset experiments reveal that msSE-ResNet with multiscale channel recalibration outperforms SE-ResNet with a single feature scale, as well as the network framework ResNet and the model with spatial and channel recalibration. scSE-ResNet results. The experimental findings on breast cancer pathology pictures at various magnifications demonstrate that the developed msSE-ResNet can be utilized for different magnifications since both the training set and the test set of the network include breast cancer pathology images at different magnifications. Breast pathological pictures with multiples have strong resilience and may be used to classify breast cancer pathological images more effectively. Furthermore, further study is needed into the selection of the compression ratio in the channel recalibration model, as well as the link between multiscale channel recalibration and classification accuracy for convolutional layers at various places in the convolutional neural network.

Various computer vision and machine learning algorithms have been employed for assessing pathological pictures at a microscopic precision as a result of the development of digital imaging methods over the past ten years. These methods could assist in automating some of the problematic workflow-related duties in the diagnostic system. For application in clinical settings, a reliable and effective image processing method is required. Regrettably, conventional methods fall short of expectations. As a result, we are still a long way from using automated breast cancer screening based on histological pictures in clinical settings. These methods, despite their great success in medical imaging, require a lot of label data, which is still lacking in this field of applications for a variety of reasons. Most importantly, annotating a dataset is quite expensive and needs a great deal of knowledge. Future research can focus on issues such as cell overlapping and uneven color distribution in pathological pictures of breast cancer created using various staining techniques.

Data Availability

The data can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] X. Feng, L. Song, S. Wang et al., "Accurate prediction of neoadjuvant chemotherapy pathological complete remission (pCR) for the four sub-types of breast cancer," *IEEE Access*, vol. 7, Article ID 134697, 2019.

[2] M. Xuru, "A classification method of breast pathological image based on residual learning," in *Proceedings of the 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, pp. 135–139, Chongqing, China, July 2020.

[3] R. Yan, F. Ren, Z. Wang et al., "A hybrid convolutional and recurrent deep neural network for breast cancer pathological image classification," in *Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 957–962, Madrid, Spain, 2018.

[4] S. Angara, M. Robinson, and P. Guillén-Rondon, "Convolutional neural networks for breast cancer histopathological image classification," in *Proceedings of the 2018 4th International Conference on Big Data and Information Analytics (BigDIA)*, pp. 1–6, Houston, TX, USA, 2018.

[5] Z. Senousy, M. M. Abdelsamea, M. M. Gaber et al., "MCUa: multi-level context and uncertainty aware dynamic deep ensemble for breast cancer histology image classification," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 2, pp. 818–829, 2022.

[6] R. Yan, J. Li, X. Rao et al., "NANet: nuclei-aware network for grading of breast cancer in HE stained pathological images," in *Proceedings of the 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 865–870, Seoul, Korea, 2020.

[7] B. Wei, Z. Han, X. He, and Y. Yin, "Deep Learning Model Based Breast Cancer Histopathological Image Classification," in *Proceedings of the 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, pp. 348–353, London, UK, 2017.

[8] Y. Liang, J. Yang, X. Quan, and H. Zhang, "Metastatic breast cancer recognition in histopathology images using convolutional neural network with attention mechanism," in *Proceedings of the 2019 Chinese Automation Congress (CAC)*, pp. 2922–2926, Hangzhou, China, 2019.

[9] S. He, R. Jun, L. Yi, W. Jianlian, W. Chenchen, and Y. Guanglu, "Combining deep learning with traditional features for classification and segmentation of pathological images of breast cancer," in *Proceedings of the 2018 11th International Symposium on Computational Intelligence and Design (ISCID)*, pp. 3–6, Hangzhou, China, 2018.

[10] G. Ye, R. Jun, W. Chenchen, Z. Jingfan, H. Simin, and W. Jianlian, "Multitask classification of breast cancer pathological images using SE-DenseNet," in *Proceedings of the 2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI)*, pp. 173–178, Guilin, China, 2019.

[11] M. Yang, P. Kumar, J. Bhola, and M. Shabaz, "Development of image recognition software based on artificial intelligence algorithm for the efficient sorting of apple fruit," in

*International Journal of System Assurance Engineering and Management*Springer Science and Business Media LLC, Berlin, Germany, 2021.

[12] J. Yan and B. Wang, "Two and multiple categorization of breast pathological images by transfer learning," in *Proceedings of the 2021 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, pp. 84–88, Oita, Japan, 2021.

[13] M. Sharma and A. Gupta, "Intercloud resource discovery: a future perspective using blockchain technology," *Journal of Technology Management for Growing Economies*, vol. 10, no. 2, pp. 89–96, 2019.

[14] H. Su, "Robust automatic breast cancer staging using a combination of functional genomics and image-omics," in *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7226–7229, Milano, Italy, 2015.

[15] V. Jagota, M. Luthra, J. Bhola, A. Sharma, and M. Shabaz, "A secure energy-aware game theory (SEGaT) mechanism for coordination in WSANs," *International Journal of Swarm Intelligence Research*, vol. 13, no. 2, pp. 1–16, 2022.

[16] T. Jakubowska, B. Wiecek, M. Wysocki, C. Drews-Peszynski, and M. Strzelecki, "Classification of breast thermal images using artificial neural networks," in *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1155–1158, San Francisco, CA, USA, 2004.

[17] S. Fan, R. Xu, and Z. Yan, "A medical pre-diagnosis system for histopathological image of breast cancer," in *Proceedings of the 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp. 1–7, Shanghai, China, 2021.

[18] C. Li, H. Niu, M. Shabaz, and K. Kajal, "Design and implementation of intelligent monitoring system for platform security gate based on wireless communication technology using ML," *Int J Syst Assur Eng Manag*, vol. 13, no. S1, pp. 298–304, 2022.

[19] M. A. R. Nayeem, M. A. M. Joadder, S. A. Shetu, F. R. Jamil, and A. A. Helal, "Feature selection for breast cancer detection from ultrasound images," in *Proceedings of the 2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, pp. 1–6, Dhaka, Bangladesh, 2014.

[20] X. Zhang, D. He, Y. Zheng et al., "Deep learning based analysis of breast cancer using advanced ensemble classifier and linear discriminant analysis," *IEEE Access*, vol. 8, Article ID 120208, 2020.

[21] P. D. Moyya, M. Asaithambi, and A. K. Ramaniharan, "Extraction of radiomic features from breast DCE-MRI responds to pathological changes in patients during neoadjuvant chemotherapy treatment," in *Proceedings of the 2019 IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 1–5, Abu Dhabi, UAE, 2019.

[22] S. R. Qureshi and A. Gupta, "Towards Efficient Big Data and Data Analytics: A Review," in *Proceedings of the 2014 Conference on IT in Business, Industry and Government (CSIBIG)*, Indore, India, 2014.

[23] G. Kaur Saini, H. Chouhan, S. Kori et al., "Recognition of human sentiment from image using machine learning," *Annals of the Romanian Society for Cell Biology*, vol. 25, 2021.

[24] S. Bagchi, M. N. H. Mohd, S. K. Debnath, M. Nafea, N. S. Suriani, and Y. Nizam, "Performance Comparison of Pre-trained Residual Networks for Classification of the Whole Mammograms with Smaller Dataset," in *Proceedings of the 2020 IEEE Student Conference on Research and Development (SCOReD)*, pp. 368–373, Batu Pahat, Malaysia, 2020.

Research Article

Biomedical Diagnosis of Leukemia Using a Deep Learner Classifier

Tawfeeq Shawly [1] **and Ahmed A. Alsheikhy** [2]

[1]*Electrical Engineering Department, Faculty of Engineering at Rabigh, King Abdulaziz University, Jeddah, Saudi Arabia*
[2]*Electrical Engineering Department, College of Engineering, Northern Border University, Arar, Saudi Arabia*

Correspondence should be addressed to Ahmed A. Alsheikhy; aalsheikhy@nbu.edu.sa

Leukemia cancer is the most common type of cancer that occurs in childhood. The most common types are acute lymphocytic leukemia (ALL) and acute myelogenous leukemia (AML) which affect children and adults, respectively. Several health issues occur due to these cancers. Leukemia affects the bone marrow or the lymph nodes. Leukemia produces abnormal white blood cells via the bone marrow system. The affected white blood cells are unable to perform their tasks properly. Detecting leukemia usually requires taking a blood smear from a patient and working with expert hematologists who analyze the smear with a microscope. In this paper, a method to detect ALL and AML using a deep learner classifier is developed and proposed. The method detects both types, determines their severity, and creates a message that recommends next steps to patients. This approach works based on image segmentation and a convolutional neural network (CNN) tool called AlexNet. The obtained results from the proposed approach and using MATLAB reached more than 98% accuracy. The margin exists because several operations are needed to fully detect the blood cancer. A dataset of leukemia from the Kaggle site is used to test the developed method and illustrate its effectiveness. This dataset is C-NMC_Leukemia, and it consists of nearly 10 GB worth of 15,000 images. A confusion matrix of testing images is provided to prove the correctness of the presented approach. Furthermore, a comparative analysis between the proposed algorithm and some works from the literature is presented. This analysis compares the method used to extract features, the classifier that is utilized, the accuracy, the precision, and the recall. The obtained results indicate that the proposed method outperforms other works and produces better results.

1. Introduction

The World Health Organization (WHO) reported that around 19 million patients were diagnosed with cancer in 2020 [1–3]. Among them, 10 million patients died. *Cancer* is the most common cause of death since it is aggressive, and its treatments can be complicated and costly [2] [4-5]. Treatment depends on the type of the diagnosed cancer [2–4].

Leukemia is an abnormality that occurs in the blood cells [2]. Blood is responsible for delivering oxygen and nutrients to the cells inside the human body. Furthermore, it transports the produced waste from them [3, 4]. Two types of leukemia have been detected and identified: acute and chronic [2, 5–7]. The first type, acute, is the most dangerous and aggressive since it spreads rapidly, and its symptoms are more severely painful than the second type [2]. The most common types are acute lymphocytic leukemia (ALL) and acute myelogenous leukemia (AML) which affect children and adults, respectively. ALL influences the white blood cells [2, 6–8]. This effect creates an unnecessary number of variations on the white blood cells. ALL occurs in children of age between 3 and 7 [2], and nearly two-thirds of diagnosed reported cases have occurred before age of 6 [2]. WHO claims that leukemia is the primary source of high death rates from cancer [3]. Chronic leukemia grows slower than acute leukemia. Acute leukemia occurs when most of the cells cannot perform their functions, whereas chronic leukemia happens when the normal cells perform their duties while some cells are immature. This situation becomes more threatening as time passes, but there is also a longer window for detection. In cases of acute leukemia, quick detection is critically important.

Blood is composed of red blood cells, white blood cells, and platelets. The red cells carry oxygen to supply the entire organ system, while the white cells protect the body from infections that occur from viruses or bacteria. The platelets support the blood clot process. When leukemia occurs, the body generates more white cells that affect other organs.

These blood abnormalities are detected via a blood smear or sample analyzed with a microscopic machine. Hematologists are key in identifying and classifying leukemia since this process depends on their experience [2, 7–10]. Numerous methods and technologies have been developed and proposed to assist those physicians in detecting and classifying both forms of leukemia. Among these approaches, pattern recognition is widely used in computer-aided systems along with image processing [2]. The time-consuming morphological process is critical in classifying blood cells, and undertaking it requires special skills. Figure 1 illustrates normal blood cells versus leukemia blood cells.

Segmentation of blood cells is vital to indicating whether blood is healthy or not [5]. Leukemia refers to diseases where some cells are divided without control and cause harm to other tissues [5, 8–11]. Hematologists focus on white blood cells since numerous infections are distinguished by those cells [5–12]. Hematologists classify leukemia as the unusual development of white platelets as depicted in Figure 1 [5]. Diagnosis of leukemia is performed as particular symptoms and signs develop in a patient, which may include frequent infections, losing unplanned weight, and weakness [5]. In addition, fever, pain in the bones, vomiting, and night sweats are symptoms of leukemia, and patients need to pay attention if one or more of these signs occur.

Acute lymphocytic leukemia (ALL), acute myelogenous leukemia (AML), chronic lymphocytic leukemia (CLL), and chronic myelogenous leukemia (CML) are the main four types of leukemia determined and identified by physicians and researchers. ALL starts in the bone marrow, and it is the most common type in children. AML is the most common type in adults. Since ALL and AML are the most common types, this research herein focuses on implementing a fast and reliable algorithm to detect them accurately and precisely.

Researchers and physical physicians have tried to determine the real cause of leukemia with no luck. Numerous factors can trigger leukemia, including radiation exposure and family history of the disease.

Physicians and other healthcare providers can be notified about a possible diagnosis of leukemia by the results of routine blood tests, but additional procedures may be required to have an accurate diagnosis including physical exams, complete blood count (CBC) tests, spinal taps, bone marrow biopsies, and imaging tests, such as CTs and MRIs. Once a patient is diagnosed with leukemia, then their healthcare providers or physicians perform treatments. These treatments may include chemotherapy, radiation, surgery, and biological therapy.

Researchers have developed and proposed numerous systems and approaches to assist physicians and hematologists in diagnosis to achieve higher accuracy [5, 11, 14]. These systems can be used to speed up the diagnosis of

leukemia [6]. The highly efficient method to diagnose leukemia uses convolutional neural networks (CNNs) [5–7]. CNNs are difficult to deploy since they are associated with a high computational cost [19, 20].

This paper proposes a method to detect AAL and AML efficiently. It develops and proposes a feasible and reliable method to detect leukemia in real time while maintaining high accuracy. This approach detects AAL and AML based on the convolutional neural network (CNN) and image segmentation using MATLAB.

The remainder of this paper is organized as follows: a literature review is presented in Section 2, and Section 3 provides details about the developing approach. Discussion and results are provided in Section 4, and the conclusion is given in Section 5.

Mondal et al. in [1] used CNNs to automate the detection of AAL from microscopic images. The authors recommended a classifier based on the weighted ensemble of different deep CNNs. Accuracy, F1-score, and kappa values were the performance metrics that were evaluated according to the weighted ensemble method. The obtained accuracy was nearly 86%, and the F1-score was approximately 89%. In this paper, for comparison, the achieved accuracy is nearly 94% and it can detect both types of leukemia at an early stage.

Oliveira and Dantas in [2] proposed a simple alteration to standard neural network (NN) construction to reach higher performance in the classification problem of the malignant leukocyte. Three constructions were tested to verify the proposed approach. In addition, around 93% of the F1-score was achieved when tested on the three constructions. Several metrics were evaluated, namely, accuracy, precision, sensitivity, specificity, and F1-score. In this proposed method, accuracy, the number of defected blood cells, and the percentage of cancer are the metrics that are considered. The developed approach can also detect ALL and AML with accuracy over 97%. Interested readers can find more information in [2].

In [3], Shaheen et al. developed a model to classify and detect AML in microscopic images based on the AlexNet approach. The authors claimed that their model reached 89% accuracy and nearly 88% precision on a dataset that contained 4 thousand blood smears. A comparison study between AlexNet and LeNet was conducted, and it showed that there was a slight difference between both models in which AlexNet performed better than LeNet. More information can be found in [3]. Herein, the proposed approach reaches almost 98% of accuracy for AAL and AML, while the method in [3] could only identify AML.

Sashank et al. in [6] proposed two different classification methods to detect AAL using deep learning techniques. An ALL-IDB2 dataset was utilized, and it contained microscopic images of blood samples. The authors used AlexNet and a machine learning model to detect ALL. CNNs, SVM, KNN, XGBoost, and decision tree were utilized as well. The obtained results from the second approach were better than those from the first approach, and the highest obtained accuracy was 100% in classification as reported by the authors. The used dataset contained 760 lymphocyte images,
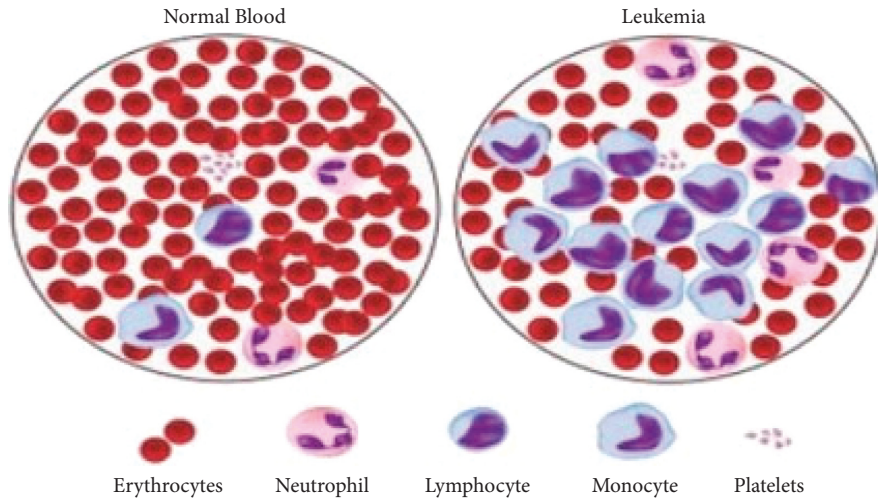
FIGURE 1: Normal blood vs leukemia blood cells.
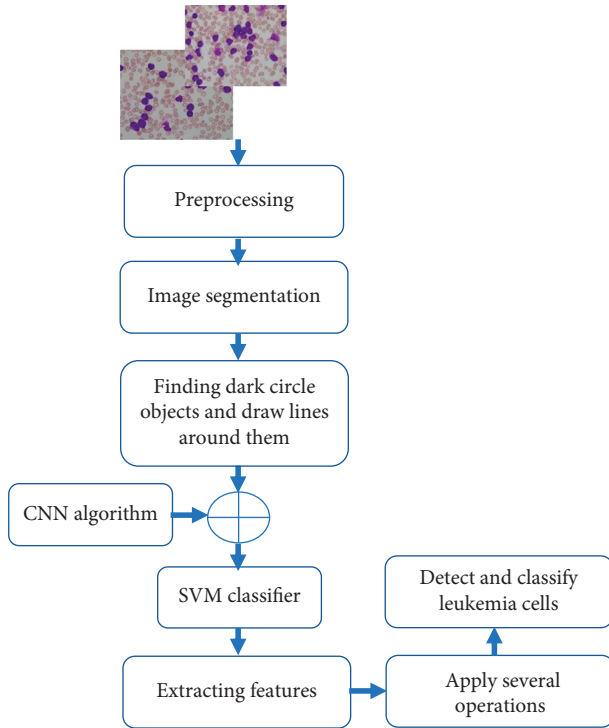
FIGURE 2: Flowchart of the proposed algorithm.
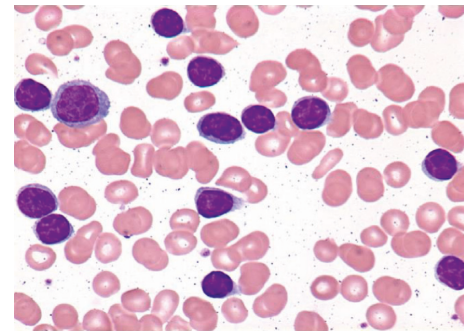
FIGURE 3: The original sample of AAL.
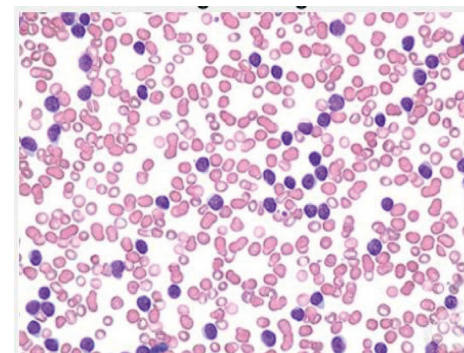
FIGURE 4: The original sample of AML.

and from these, 570 images were used for training while the rest were used for testing. Readers can get more information in [6].

Claro et al. in [7] presented a CNN architecture to differentiate blood slides that contained ALL, AML, and healthy blood slides (HBS). In [7], 16 datasets were utilized to conduct several experiments. These datasets contained 2415 images, and the method obtained 97% accuracy and precision. The authors performed a comparison experiment with numerous methods that used CNNs only. More information in [7] can be obtained for interested readers.

In [9], Dasariraju et al. presented a method to detect and classify AML using a machine learning model based on analyzing immature leukocytes. The authors obtained their dataset from the *Cancer* Imaging Archive, which contained data from AML patients and healthy patients. The authors used image format conversion, multi-Otsu threshold, and morphological process. In addition, 16 features were extracted from every image. A random forest algorithm (RFA) was used to train the dataset, and it produced nearly 93% of accuracy in detection and almost 94% in
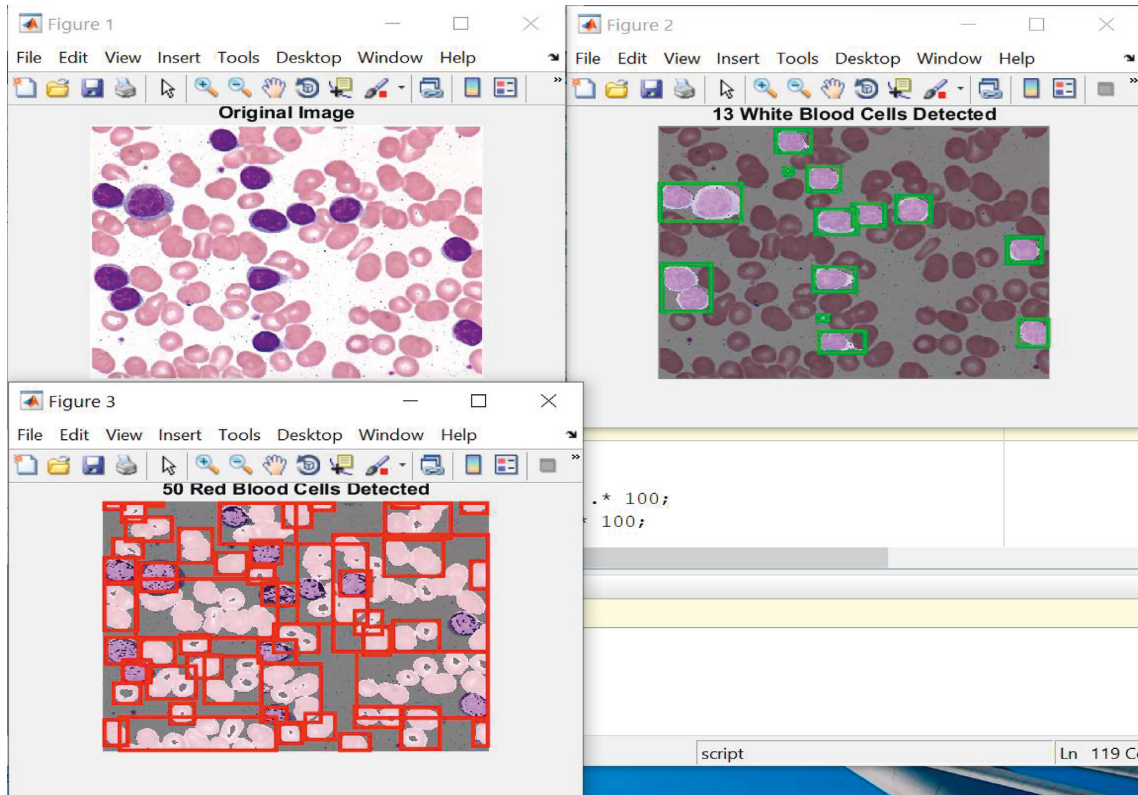
FIGURE 5: Sample outputs from the developed method.
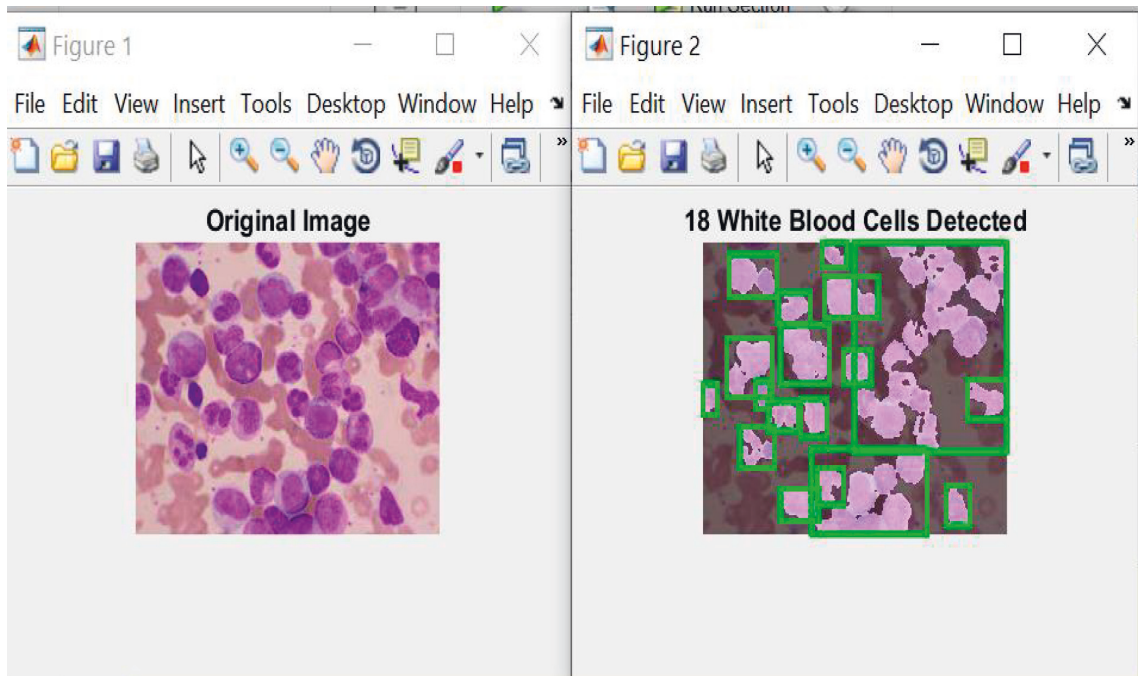
FIGURE 6: Detected white blood cells.

classification. However, the obtained precision only reached 65%, far less than the model in this paper. The developed model in this paper also detects AAL and AML with an accuracy over of 98% exceeding the presented model in [9].

Pallegama et al. in [10] proposed a method to detect ALL cells using CNNs. The authors claimed that their approach could reduce the time needed for analyzing the blood samples and the cost for microscopic observations. Over 100
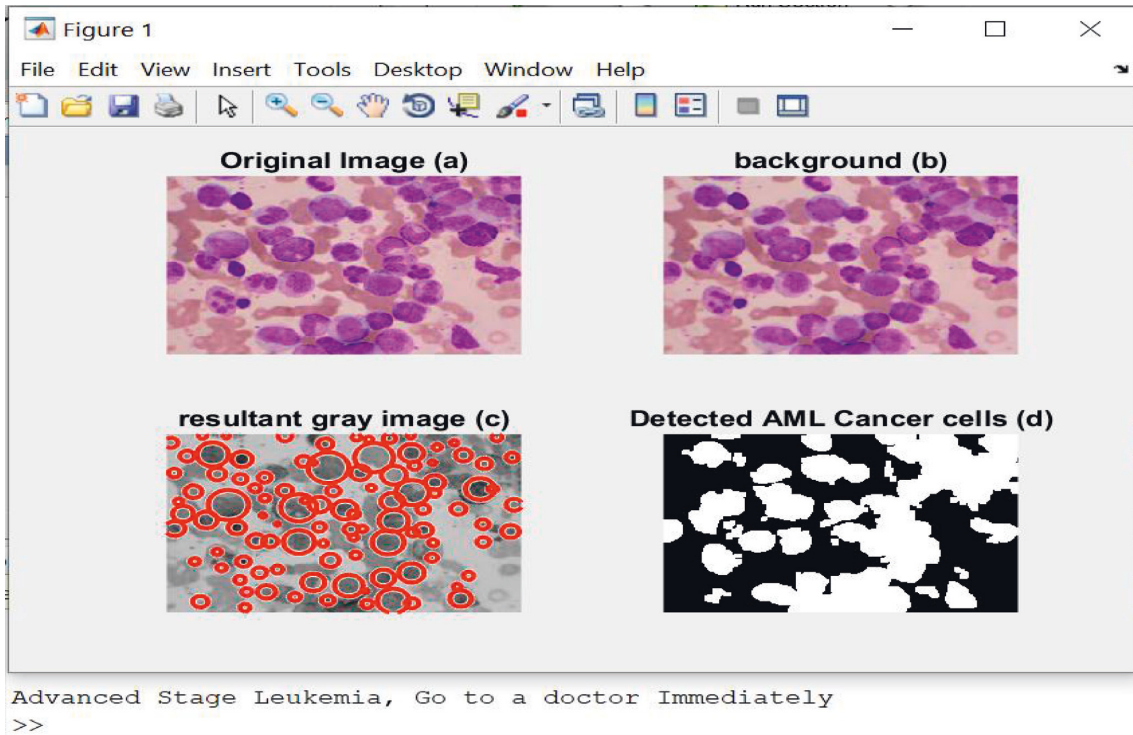
FIGURE 7: The obtained outputs of AML leukemia.
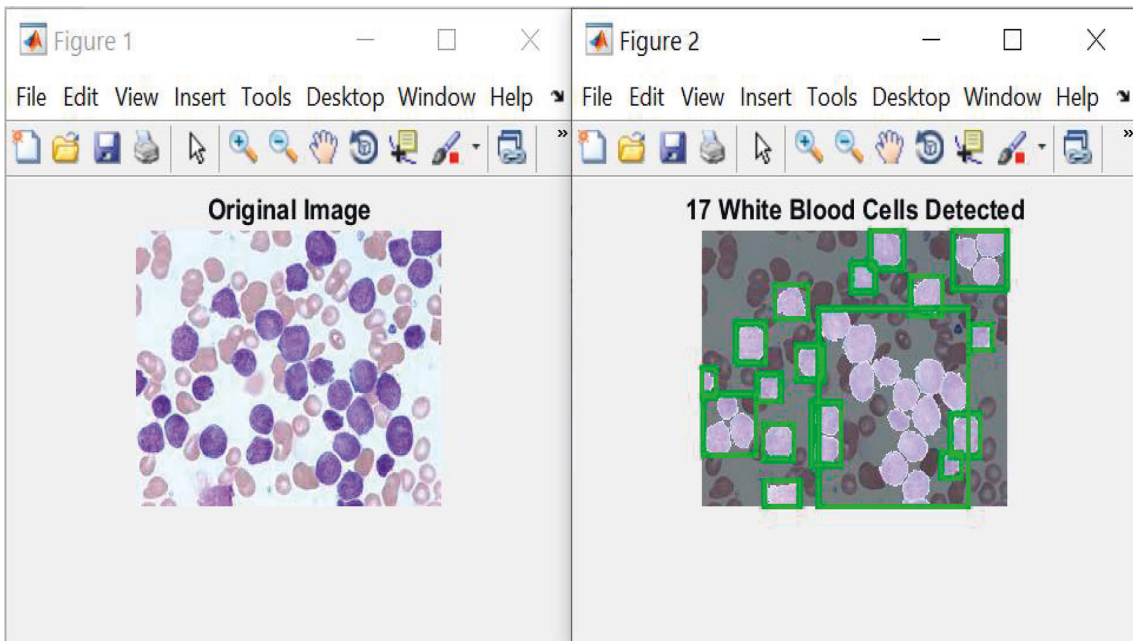
FIGURE 8: Obtained results for detecting white blood cells affected by ALL.

blood smears were used to train the method to detect ALL cells. These blood samples were diagnosed ALL by a cancer hospital. The proposed approach herein detects ALL and AML, so it is better than method that was presented in [10].

Loey et al. in [11] presented two automation methods based on a transfer learning approach to detect leukemia. In the first method, a pretrained AlexNet was used to extract features from blood microscopic images. In the second method, fine-tuning was performed for all extracted features to detect leukemia. Both methods were tested on a dataset that contained around 3000 images. The second approach performed better than the first one in classification and claims 100% accuracy. Interested readers can refer to [11] for more information.
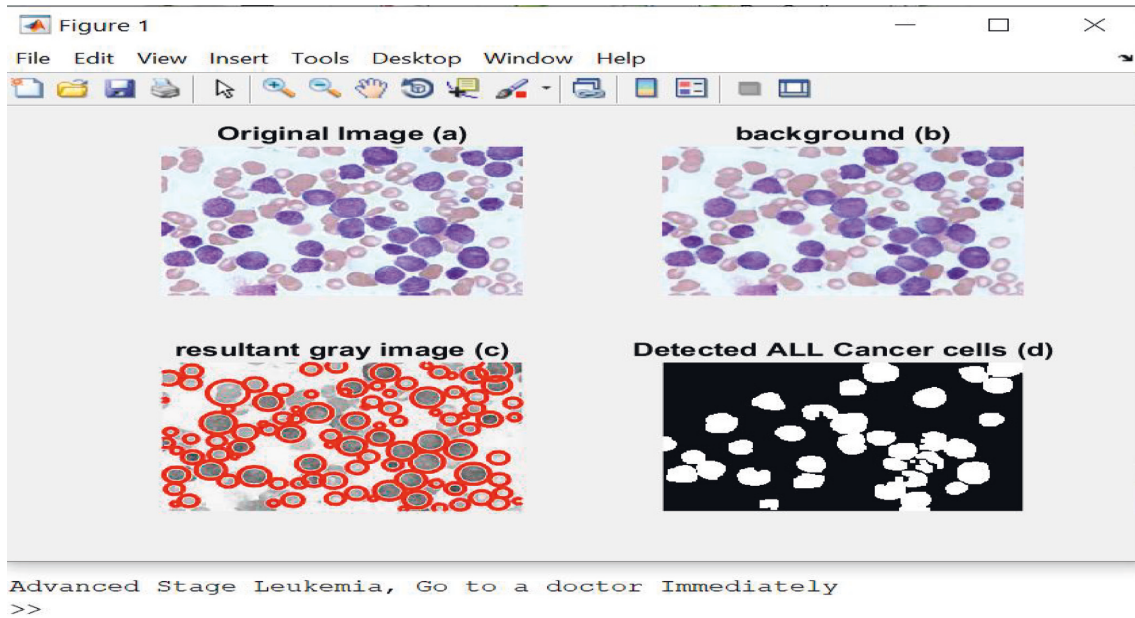
Figure 9: The obtained results of ALL leukemia detection.

Bhandari et al. in [21] performed a comprehensive analysis of the state-of-the-art methods to detect cancer utilizing genetic algorithms. The authors made a deep analysis to identify the future challenges in the development of such techniques. This analysis was related to various types of cancer such as bladder, breast, ovarian, and leukemia. The authors focused on the type of cancer, functions being used, the main purpose of the methods, and the type of data being tested and verified. Additional information can be in [21].

In [22], Hamza et al. implemented a method to detect and classify ALL using an optimal deep transfer learning method. Blood smear images were utilized for detection and classification purposes. A filter was used to remove noise, and the fuzzy c-means method was involved to segment the inputs. Features were extracted using the competitive swarm optimization and NetB0 approaches. The authors measured several performance metrics including accuracy, precision, recall, specificity, and F-score. The authors claimed that their algorithm achieved 96%, 95.715%, and 96.51% accuracy, precision, and recall, respectively. Even so, our approach detects and classifies ALL and AML with higher accuracy, precision, and recall. These results indicate that the proposed algorithm herein outperforms the implemented method in [22].

Abir et al. in [23] developed a method to detect ALL using a transfer learning model. This method achieved nearly 98.3% of accuracy, while our presented algorithm achieves nearly over 99% of accuracy. Four different types of models were utilized. However, these four models detected only ALL, while our algorithm detects and classifies ALL and AML as well. Additional information is found in [23].

2. Materials and Methods

The proposed approach began with a patient who suffered greatly from leukemia. He was diagnosed with leukemia when he was 6 years old. Initially, his physician diagnosed him with an infection, and the leukemia was only detected later. Due to his advanced condition, he had to go through a complex treatment, but the cost was too high. It became critical to consider a new method for leukemia detection. The proposed method detects ALL and AML since both types occur more than others. This paper uses an 8-layer CNN called AlexNet. All images used are 227×227 pixels in size. MATLAB is used as a programming platform and a simulation tool to train the developed approach. The proposed algorithm contains several processes as depicted in Figure 2.

The proposed approach is illustrated in the following Algorithm 1:

AlexNet is involved to extract features of the white blood cells to determine healthy and infected cells. The presented algorithm learns itself regularly according to the obtained results. The extracted features include mean squared error (MSE), histogram of oriented gradients (HOG), and local binary pattern (LBP). In addition, other features are extracted and utilized as well.

One dataset was used to train, validate, and test the proposed algorithm. In addition, some metrics are evaluated during the simulation including the following:

(1) True Positive (TP): it measures a total number of correctly identified blood samples to detect leukemia ALL and AML.

(2) False Positive (FP): it refers to incorrectly identified samples.

(3) True Negative (TN): it defines a number of total negative samples that were detected and classified correctly by the algorithm.

(4) False Negative (FN): it provides an indication of a total number of the negative samples that were incorrectly classified and detected.

(i) Input: an image to read.
(ii) Output: the detection and classification of leukemia: AAL or AML.
(1) **Read an image from a file and display it**.
(2) **In the preprocessing phase: separate foreground and background**.
(3) **Transform the resultant image into gray image**.
(4) **Extract values of RGB from the original image**.
(5) Image segmentation: mapping between foreground values and RGB values is performed to increase the contrast.
(6) Compute the radius of every blood cell in the corresponding gray image and save results in a matrix radi.
(7) Determine black cells and estimate their radius as well.
(8) Draw a red line around every dark cell, determine their numbers, and save results in a variable *x*.
(9) Detect white blood cells and draw green rectangle around them.
(10) Determine number of white and red blood cells.
(11) Calculate a threshold of every detected infected cell using Otsu's approach to minimize the variance between the white and black cells in the gray image and save the results in a variable thre.
(12) Convert the gray image into the binary image using the threshold to locate the potential areas of all detected infected cells.
(13) Label all infected cells and mark them on the original image as well.
(14) Remove all unwanted, healthy blood cells using the erosion function.
(15) Segment all contiguous regions of interest into distinct objects using a built-in function.
(16) Determine the location of infected cells and their corners using the Harris–Stephens method.
(17) Extract features from infected cells using AlexNet, CNN, technique.
(18) **For i = 1: number of infected blood cells**.
(19) Apply SVM to classify every infected blood cell whether it is AAL or AML.
(20) Determine the percentage of leukemia and save result in a variable *y*.
(21) Determine the status of leukemia.
(22) Display a message to patients.
(23) **End**
(24) **End of algorithm**.

ALGORITHM 1: Leukemia detection and classification.

TABLE 1: Evaluated performance metrics.

Number of testing images: 2,500.	
Metric	Evaluated value
TP	1981
FP	14
TN	494
FN	11
PRE	99.30%
REC	99.45%
Accuracy	99%

TABLE 2: Confusion matrix of the testing dataset.

		True class	
Predicted Class		ALL	AML
	ALL	953 = (98.55%)	14 = (1.45%)
	AML	11 = (1.08)	1002 = (98.92%)

that are detected and classified correctly over the total number of samples as depicted in

$$Accuracy = \frac{(TP + FN)}{(TP + TN + FN + FP)}. \quad (3)$$

As stated earlier, the proposed method uses one dataset in which around 10,500 blood samples, 70% of the dataset, are assigned for the training purpose. The remaining 30% of the dataset is divided into two groups: 15% for testing and 15% for validating the results. For the validation, there are 2,250 images of blood samples. Figures 3 and 4 illustrate the original images of ALL and AML, respectively. In the training stage, the samples are either healthy or infected. For every input or sequence of inputs, the presented algorithm extracts feature from healthy samples and infected samples as well. These features are deeply analyzed in order for the implemented model to be able to determine and classify ALL and AML accurately.

Figure 5 illustrates the obtained outputs from the presented approach in which white and red blood cells detected in a random blood sample are surrounded by green and red

(5) Precision (PRE): this metric measures a fraction of the true samples that are identified correctly over the same samples as well as the samples that are classified incorrectly as shown in

$$PRE = \frac{TP}{(TP + FP)}. \quad (1)$$

(6) Recall (REC): it indicates the fraction of the truly identified samples over the summation of true samples plus the number of negative samples that are classified incorrectly as depicted in

$$REC = \frac{TP}{(TP + FN)}. \quad (2)$$

(7) Accuracy: this metric shows the percentage of the summation of the true samples and the negative ones

Table 3: Comparison of three-performance metrics.

Developed method	Method to extract features	Classifier name	Accuracy	Precision	Recall (%)
Mondal et al., 2021 [1]	CNN	CNN	86.2%	88.7%	88.8
Oliveira and Dantas, 2021 [2]	CNN	CNN	91.49%	89.61%	93.90
Shaheen et al., 2021 [3]	CNN	AlexNet	98.58%	87.4%	88.9
Sashank et al., 2021 [6]	CNN	SVM, KNN, decision trees	95.05%	95.25	96.75
Claro et al., 2020 [7]	CNN	CNN	97.18%	97.23%	97.18
Dasariraji et al., 2020 [9]	Random forest	Random forest	92.99%	91.23%	95.41
Loey et al., 2020 [11]	CNN	FC	99.04	99.64%	98.44
The proposed method	CNN (AlexNet) + SVM	SVM	99.30%	99.45%	99

rectangles, respectively. The upper left image shows the original image, the upper right picture represents the resultant image of the detected white cells, while the bottom left picture denotes the detected red blood cells of the sample image.

The proposed algorithm identifies the white and the red blood cells in the input images as shown in Figure 5. The white blood cells are counted and surrounded by green rectangles, while red rectangles are drawn around every red cell, as illustrated in Figure 5. These results are utilized later in the deep learning and classification phases.

3. Results and Discussion

MATLAB is used to conduct several experiments to process images of blood samples that are either healthy or infected. The infected samples are infected by ALL and AML. To detect and classify AML or ALL using the presented algorithm, 10,500 images of blood smears are used to train the algorithm in detection and classification, while 2,500 images are utilized for validation. The remaining images are used and utilized for testing. Support vector machine (SVM) performs the classification operation. The 2,500 of testing images are also used to evaluate accuracy, precision, and recall.

Example 1. Leukemia: AML cells.
Figure 6 displays the detected white blood cells of AML type. These cells are surrounded by the green rectangles. Figure 7 illustrates the obtained results of the detection and classification operations.

Figure 7(a) represents the original image, 7(b) shows the original image after removing noises, while 7(d) highlights the detected AML cancerous cells in white color.

Example 2. Leukemia: ALL cells.
Figure 8 shows the detected white blood cells of ALL enclosed by the green rectangles, and Figure 9 demonstrates the obtained results along with a message to the patient.

Table 1 lists the values of all mentioned metrics that were determined by the developed approach.

The proposed approach detects both types of leukemia with an accuracy of over 98% as shown in Table 1. Table 2 depicts the confusion results on the testing dataset represented in a confusion matrix. The corrected identified results are distinguished in green, while red boxes refer to the inaccurate detection and classification of ALL and AML, respectively.

The comparison study between the presented algorithm and other developed approaches in the literature is conducted and shown in Table 3. This comparison evaluates three-performance metrics which are precision, recall, and accuracy. All results are given in Table 3. This table lists the works that were developed with their references' numbers, the methodologies used to extract features, their classifiers, and the three-performance metrics for comparison. The values for all metrics are the overall obtained results.

Table 3 shows that the proposed algorithm herein outperforms and outstands most of the developed and implemented approaches in the literature. This indicates that the developed approach in this research produces promising results.

4. Conclusions

The proposed method has the capability to detect and classify ALL and AML cancer with high precision and accuracy as proved by the conducted experiments. Hence, it can be used in hospitals and healthcare centers to support and assist hematologists and laboratory technicians in their tasks. In addition, the developed algorithm reaches an accuracy of nearly 99% in detection and classification.

Data Availability

The authors would like to confirm that the dataset which is utilized in this research is available at the Kaggle website and can be found at the following link: https://www.kaggle.com/datasets/andrewmvd/leukemia-classification.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] C. Mondal, K. Hasan, T. Jawad et al., "Acute Lymphoblastic Leukemia Detection from Microscopic Images Using Weighted Ensemble of Convolutional Neural Networks," pp. 1–31, 2021, https://arxiv.org/abs/2105.03995.

[2] J. E. M. D. Oliveira and D. O. Dantas, "Classification of Normal versus Leukemic Cells with Data Augmentation and Convolutional Neural Networks," in *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, vol. 4, pp. 685–692, January 2021.

[3] M. Shaheen, R. Khan, R. R. Biswal et al., "Acute myeloid leukemia (AML) detection using AlexNet model," *Complexity*, vol. 2021, Article ID 6658192, 8 pages, 2021.

[4] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajjej, and A. A. Shihab, "Applying dynamic systems to social media by using controlling stability," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4569879, 7 pages, 2022.

[5] N. Boykoff, L. Freage, J. Lenn, and P. Mallikaratchy, "Bispecific aptamer sensor toward T-cell leukemia detection in the tumor microenvironment," *ACS Omega*, vol. 6, no. 48, Article ID 32563, 2021.

[6] G. V. S. Sashank, C. Jain, and N. Venkateswaran, "Detection of acute lymphoblastic leukemia by utilizing deep learning methods," in *Machine Vision and Augmented Intelligence—Theory and Applications*, M. K. Bajpai, K. Kumar Singh, and G. Giakos, Eds., Springer, Berlin/Heidelberg, Germany, 2021.

[7] M. Claro, L. Vogado, R. Veras et al., "Convolution Neural Network Models for Acute Leukemia Diagnosis," in *Proceedings of the International Conference on Systems, Signals and Image Processing*, pp. 63–68, IWSSIP, Niteroi, Brazil, July 2020.

[8] M. B. Alazzam, A. T. Al-Radaideh, N. Binsaif, A. S. AlGhamdi, and M. A. Rahman, "Advanced deep learning human herpes virus 6 (HHV-6) molecular detection in understanding human infertility," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1422963, 5 pages, 2022.

[9] S. Dasariraju, M. Huo, and S. McCalla, "Detection and classification of immature leukocytes for diagnosis of acute myeloid leukemia using random forest algorithm," *Bioengineering*, vol. 7, no. 4, pp. 120–212, 2020.

[10] R. D. A. U. Pallegama, B. G. D. A. Madhusanka, and H. D. N. S. Priyankara, "Acute lymphoblastic leukemia detection using convolutional neural network," *International Journal of Engineering Science and Computing (IJESC)*, vol. 10, no. 6, Article ID 26529, 2020.

[11] M. Loey, M. Naman, and H. Zayed, "Deep transfer learning in diagnosing leukemia in blood cells," *Computers*, vol. 9, no. 2, pp. 29–12, 2020.

[12] H. T. Salah, I. N. Muhsen, M. E. Salama, T. Owaidah, and S. K. Hashmi, "Machine learning applications in the diagnosis of leukemia: current trends and future directions," *The International Journal of Literary Humanities*, vol. 41, pp. 717–725, 2019.

[13] M. B. Alazzam, H. Mansour, F. Alassery, and A. Almulihi, "Machine learning implementation of a diabetic patient monitoring system using interactive E-app," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5759184, 7 pages, 2021.

[14] P. G. F. Desai and G. Shet, "Detection of leukemia using image processing," *International Journal of Advance Research in Science and Engineering (IJARSE)*, vol. 7, no. 3, pp. 149–156, 2018.

[15] S. Shafique and S. Tehsin, "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks," *Technology in Cancer Research and Treatment*, vol. 1, no. 17, Article ID 153303381880278, 2018.

[16] P. V. R. Raju, G. R. Babu, G. N. Raji, G. Pavani, K. Srikanth, and J. C. L. Sobha, "Detection of leukemia using image processing," *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 4, no. 11, pp. 123–127, 2017.

[17] M. B. Alazzam, A. T. Al-Radaideh, R. A. Alhamarnah, F. Alassery, F. Hajjej, and A. Halasa, "A survey research on the willingness of gynecologists to employ mobile health applications," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1220374, 7 pages, 2021.

[18] SA. Rejintal and N. Aswini, "Image processing based leukemia cancer cell detection, international journal of engineering research and technology (IJERT)," in *Proceedings of the ICACT Conference*, PyeongChang, Korea (South), November, 2016.

[19] H. P. Vaghela, H. Modi, M. Pandya, and M. B. Potdar, "Leukemia detection using digital image processing techniques," *International Journal of Applied Information Systems*, vol. 10, no. 1, pp. 43–51, 2015.

[20] N. Patel and A. Mishra, "Automated Leukemia Detection Using Microscopic Images," *Second International Symposium on Computer Vision and the Internet*, vol. 58, pp. 635–642, 2015.

[21] A. Bhandari, B. K. Tripathy, K. Jawad, S. Bhatia, M. K. I. Rahmani, and A. Mashat, "Cancer detection and prediction using genetic algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, Article ID 1871841, 18 pages, 2022.

[22] M. A. Hamza, A. A. Albraikan, J. S. Alzahrani et al., "Optimal deep transfer learning-based human-centric biomedical diagnosis for acute lymphoblastic leukemia detection," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, Article ID 7954111, 13 pages, 2022.

[23] W. H. Abir, F. Uddin, F. R. Khanam et al., "Explainable AI in diagnosing and anticipating leukemia using transfer learning method," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5140148, 14 pages, 2022.

[24] Z. Jiang, Z. Dong, L. Wang, and W. Jiang, "Methods for Diagnosis of Acute Lymphoblastic Leukemia Based on ViT-CNN Ensemble Model," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7529893, 12 pages, 2021.

[25] S. Rezayi, N. Mohammadzadeh, H. Bouraghi, S. Saeedi, and A. Mohammadpour, "Timely diagnosis of acute lymphoblastic leukemia using artificial intelligence-oriented deep learning methods," *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, Article ID 5478157, 12 pages, 2021.

Research Article

The Use of Thinking Visualization Techniques in College Teaching Based on Improved Genetic Algorithms

Luyan Su (ID)

Chengyi University College, Jimei University, Xiamen, Fujian 361021, China

Correspondence should be addressed to Luyan Su; suluyan@jmu.edu.cn

Current educational resources do not maximize energy efficiency, and scientific and proven teaching methods are necessary for today's university education to help achieve the integration of teaching resources and improve teaching quality. This study discusses the application of artificial intelligence and thinking visualization technology in college education and teaching, firstly introduces the advantageous role of genetic algorithm in artificial intelligence and the energy efficiency of thinking visualization technology, then conducts a research analysis, introduces genetic algorithm as technical support in teaching and learning, uses thinking visualization technology as a tool, aims to study the application of thinking visualization technology in college teaching, and expects to use genetic algorithm and thinking visualization techniques to improve teaching effectiveness. The experimental results show that genetic algorithm and thinking visualization technology can effectively optimize the co-giving of educational services, reduce the workload of teachers, while reducing the cost of learning, enhance the learning effect and experience, and improve students' efficient thinking.

1. Introduction

With the continuous improvement of the economic level, the curriculum is constantly changing in order to adapt to the current development, from the initial subject education to quality education. In the development process of the overall quality education, vocal music occupies an important position, and the educational nature of music is becoming more and more prominent. Vocal music teaching is the main content of music teaching. With the continuous exchange of Chinese and foreign cultures, we found that there are many excellent musical skills that deserve attention. Looking at the current music market, there are many outstanding music representatives, but there are still many deficiencies. For example, when teachers are teaching, they focus on singing songs, ignoring the understanding of the lyrics and songs themselves. Their grasp of lyrics and music remains on the surface, and they cannot grasp the core things. This study aims to study the application of thinking visualization technology in vocal music teaching by integrating and improving genetic algorithm, and expects to use genetic algorithm and thinking visualization technology to improve the effect of vocal music teaching.

In the development of the 21st century, we are no longer limited only by the economy. Vocal music education also adds color to the world. Vocal music teaching has far-reaching significance for the development of music education in the future. It is necessary to change the traditional concept of education, strengthen the comprehensive quality education of students, and provide a good employment orientation for future employment. The combination of vocal music teaching and thinking vision can improve the tacit understanding between teachers and students. This improves the overall level of vocal music teaching and singing and makes vocal music teaching systematic and scientific.

There is not a lot of intersection between the two disciplines of computing and music. But it is still a relatively new point of view in vocal music teaching. Combining the two provides a new research direction. In this study, vocal music education is intuitive and systematized, so that the learning of vocal music classroom is more abundant.

2. Related Work

Genetic algorithm has a very great application prospect at present. To understand a brief understanding of its application direction, Tavakkoli-Moghaddam et al. proposed a genetic algorithm (GA) for the redundancy assignment problem of series-parallel systems. Experiments show that genetic algorithm is an effective method to solve this kind of problem. Finally, he presents computational results for typical scenarios and discusses the robustness of the proposed algorithm [1]. Dawid and Kopel used the idea of artificial adaptive agents in economic theory. He used a genetic algorithm (GA) to simulate the learning behavior in the system. The striking differences between the results of different setups are explained. Finally, he explained the relationship between the encoding and convergence properties of GA [2]. Volkanovski proposed a new method to optimize the maintenance scheduling of power system generator sets. The maintenance plan minimizes the risk by minimizing the annual value of the expected loss of load. He also compared the results gained with those obtained by approximate methods. The results show that the maintenance schedule obtained with the new method improves the reliability of the power system compared to the results of the approximate method [3]. Long and Wu studied a global optimization method combining genetic algorithm and the Hooke–Jeeves method to solve a class of constrained optimization problems. He introduces the method of the second penalty factor and the law of precise penalty factor. When dealing with constrained optimization problems, a common approach is to add constraints as penalty terms to the objective function. "Punishment" is a very vivid name, which means that when the optimization process iterates beyond the constraints, it is punished, or negative feedback. Then, the genetic algos were combined with the Hooke–Jeeves approach to tackle the conversion followed by the post-optimization process. This methodology greatly reduces the genetic model's accuracy and its speed of convergence [4]. Yoshitomi et al. extended genetic algorithms to uncertain environments. The method was applied to the stochastic optimal assignment problem. He proved that the solution of GAUCE agrees well with the solution [5]. To solve the linear and U-shaped pipeline balance problem, Alavidoost et al. proposed an improved genetic algorithm. He selected one out of five rules of failure in the algorithm to enhance the performance of the hereditary method. Furthermore, he selected fuzzy controllers under pressure for better implementation of one out of five success rules in genetic algorithms. The outcome demonstrates the outstanding quality of his presented algorithm [6]. Metawa et al. proposed an intelligent genetic algorithm (GA)-based model. It is used to constrain bank borrowing and lending decision making. It seeks dynamic lending decisions by minimizing bank profit and also bank default probability. It allows banks to reduce their loan screening time by 12–50% [7]. Gong et al. proposed an ensemble-based genetic algorithm. He defined an ensemble-based Pareto dominance relation to modify the fast nondominated sorting method in NSGA-II. Numerical results demonstrate the superiority of the method [8].

Fetouh and Zaky proposed a new method for coordinating the stability of power systems. The proposed method is designed to minimize the computation time, reduce the storage capacity required for optimization problems, and improve the stability performance of the power system [9]. Rajeswari and Neduncheliyan have proposed a fault-tolerant technology based on clumps which uses genuine arithmetic. The genetic algorithm is used to select a set of standby knots based on sponsorship coverage and residual energy parameters. Simulation results show that his proposed technique minimizes energy loss and overhead [10]. Prichard's aim was to examine the effectiveness of music teaching among teachers enrolled in introductory music education courses. He used a sequential explanatory mixed methods design and divided it into two parts. Mixed methods analysis revealed that music education majors had higher beliefs about music teaching efficacy [11]. Steele Royston examined the significance of human connections within the context of the lesson and offered tips for proactive faculty to better their connections. Connecting with students as a musical teacher leads them to understand a whole lot with the arts of playing music. It also teaches them something about the skill of what it means to be a sympathetic and loving person [12]. The introduction of genetic algorithm and vocal music teaching in these researches is in place, but the practicability is not high due to the strong professionalism.

3. Methods

3.1. Genetic Algorithm. Genetic algorithm refers to the population of solution sets, and the population is composed of several individuals encoded by genes. In essence, a genetic algorithm is an optimization algorithm, which is a random search algorithm that uses the idea of natural selection and biological evolution to search for the optimal solution in the search space. In the process of vocal music teaching, the genetic algorithm can integrate the characteristics of students, optimize the vocal music teaching method, improve the students' vocal music mastery ability, and better serve the vocal music classroom.

Evolutionary computing is a stochastic optimization method that simulates the genetic mechanism of the animal kingdom, with the idea of "survival of the fittest." Therefore, genetic manipulation and natural selection become an important part when using evolutionary thinking to solve target optimization problems. Evolutionary algorithms iteratively generate multiple solutions to each problem during the optimization process and continue to generate better solutions. The optimization function is called the fitness function. Each solution is called an individual, and all individuals in each generation form a group. The fitness value of individuals in each group is different. With iteration, it obtains better individuals through certain evolutionary strategies, such as crossover and mutation. When the algorithm terminates, the individual with the best fitness value in the entire population is selected as the solution to the problem [13]. It is shown in Figure 1.

Evolutionary algorithms are random search methods. Compared with other enumeration techniques and heuristic
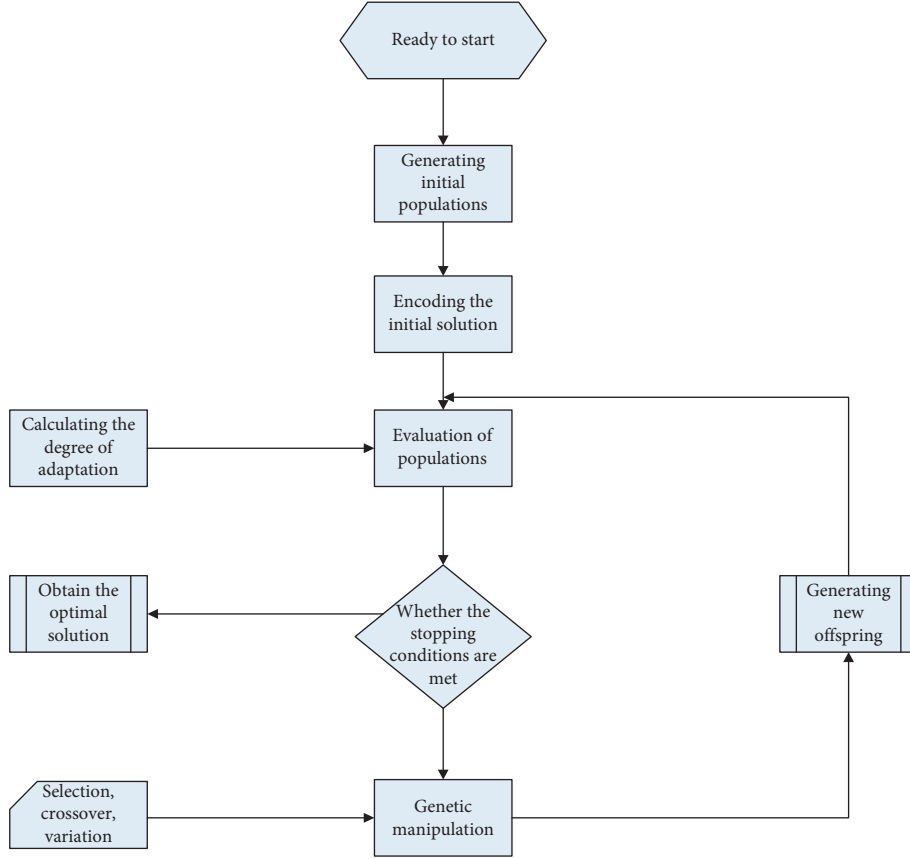
FIGURE 1: Basic framework of genetic algorithm.

search techniques, the global optimal solution probability of this problem is high. The evaluation information of the objective function is used to make it actionable and general. It can be operated with massively parallel computers and can be easily combined with other methods. At present, evolutionary computing mainly includes genetic algorithm, evolutionary strategy, and evolutionary planning. In the process of vocal music teaching, the evolutionary algorithm can be combined to find the most suitable teaching method and improve the teaching efficiency and effect.

$$\min\{g(a)|a \in Q\}. \tag{1}$$

Formula (1) represents the function expression of the function optimization model, where a represents the decision variable, $g(a)$ represents the objective function, and Q represents the spatial subset.

$$\min g(a).st^* p(a) \ge 0. \tag{2}$$

Formula (2) represents the decision variable inequality.

$$Q = \{a \in W^s | p(a) \ge 0\}. \tag{3}$$

Formula (3) represents the feasible region of the decision solution, which is the set of all solutions in layman's terms.

In order to get the optimal solution for the region, we need to optimize the model.

$$\forall a \in Q \cap \left\{a \in W^s | \sqrt{\sum (a_1 - a^*)^2} < \alpha\right\}. \tag{4}$$

When formula (4) satisfies $g(a^*) \le g(a)$, we say that the model has an optimal solution. Figure 2 shows the basic flow chart of the algorithm.

Various encoding methods have emerged today. The most commonly used genetic algorithm is binary coding. In binary, all symbols consist of "0" and "1."

$$\beta = \frac{T_{\max} - T_{\min}}{2^P - 1}. \tag{5}$$

Formula (5) represents the precision of binary encoding, where T represents the range of values and p represents the symbol length of the encoding.

$$k = T_{\min} + \left(\sum_{o=1}^{p} w_o * 2^{o-2}\right) * \frac{T_{\max} - T_{\min}}{2^P - 1}. \tag{6}$$

Formula (6) represents the decoding function expression when the encoded length is k.

$$\begin{cases} f_p = w_p, \\ f_p = w_{p+1} \oplus w_p. \end{cases} \tag{7}$$

$$\begin{cases} w_p = f_p, \\ w_p = w_{p+1} \oplus f_p. \end{cases} \tag{8}$$
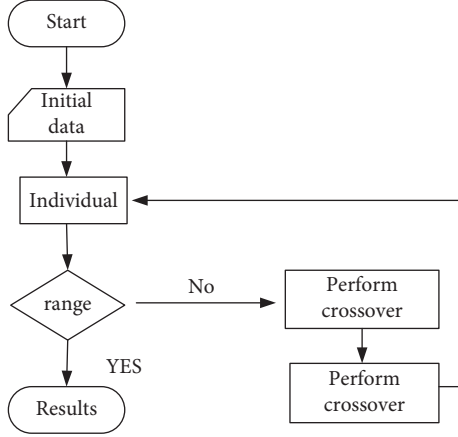
Figure 2: Model algorithm flow demonstration.

Formulas (7) and (8) represent binary codes under different Gray codes.

In the genetic algorithm, we usually use the fitness function to judge the situation of the individual. The fitness function satisfies the generality, so the calculation steps can be reduced in the actual use process.

$$\text{Fitness}\,(g\,(a)) = g\,(a).\qquad(9)$$

Formula (9) represents the functional expression of the maximal optimization problem.

$$\text{Fitness}\,(g\,(a)) = -g\,(a).\qquad(10)$$

Formula (10) represents the functional expression for the minimal optimization problem.

For multi-objective optimization problems, researchers have proposed different evolutionary multi-objective research algorithms. In the process of solving the problem, we have two goals to achieve. First, the frontier of the approximate Pareto optimal solution set should be as close as possible to the frontier of the true Pareto optimal solution set. The closeness of the two fronts reflects the convergence of the algorithm. Second, the target vectors corresponding to the Pareto optimal solution set should be distributed as widely and evenly as possible, which reflects the diversity of the knowledge set. Multi-objective optimization can be described by the following formula:

$$\begin{cases} \max G\,(i) = (g_1\,(i)g_2\,(i),\cdots,g_x\,(i))y, \\ s.y.\, i \in \theta. \end{cases}\qquad(11)$$

If the Pareto boundary resolution system is able to fulfill the condition, it means that it is practicable for the algorithm to address the MOP issue.

$$\begin{cases} \min\theta = G\,(i) \\ = g_1\,(i), g_2\,(i),\cdots, g_n\,(i)\,y \\ sy.p_t\,(i) \le 0, t = 1,2\cdots z \\ t_e\,(i) = 0, e = 1,2\cdots, q \end{cases}.\qquad(12)$$

In the formula, $i = (i_1, i_2,\cdots,i_n)$ represents the n-dimensional decision vector, and $g\,(i)$ is the objective function.

$p_t\,(i) \le 0, t = 1,2\cdots z$ defines the inequality constraint for z, and $t_e\,(i) = 0, e = 1,2\cdots, q$ defines the equality constraint function for q.

As can be seen from the above formula, the MOP problem is difficult to determine a unique solution, which is different from a single objective. In addition, the solution process of this algorithm is very complicated and difficult due to many constraints.

Standard genetic algorithms are optimization algorithms for individual objectives. The fitness function of an individual is usually the objective function of optimization. The result of the genetic algorithm is to obtain the individual with the best fitness. In reality, many problems are multi-objective problems, that is, multiple values of the problem solution need to be optimized simultaneously. For multi-objective optimization problems, a solution may be better for one objective and worse for others. The two keys to multi-objective optimization using genetic algorithms are: (1) How to carry out fitness value distribution and individual selection, so that individuals in the population converge to the nondominated solution set; (2) How to keep the population multiple and avoid premature stock convergence, so that the individuals in the population are evenly distributed on the Pareto frontier. Around these two problems, researchers have proposed different types of multi-objective evolutionary algorithms.

$$\min f^{ab}\left(y|c^i, d^*\right) = \max_{0 \le x \le 1} \frac{\left|g_x\,(a)\right| - d_x^*}{c_x^i}.\qquad(13)$$

Formula (13) represents the Chebyshev function expression, in order to avoid the case of division by zero, when $c_x = 0$, $c_x = 10^{-6}$.

$$\min f^{zxc}\left(a|d_c, k^*\right) = p_1^c + \rho p_2^c,$$
$$p_1^c = \frac{(G\,(a) - k^*)^t d^c}{\|d^c\|},\qquad(14)$$
$$p_2^c = \left\|G\,(a) - k^* - \left(\frac{p_1^c}{\|d^c\|}\right)d^c\right\|.$$

The above 3 function expressions belong to the boundary intersection method, where *Subject* to $a \in \beta$.

$$\text{Fitness}\,(g\,(a)) = \begin{cases} g\,(a) + Y_{\min}, \\ 0. \end{cases}\qquad(15)$$

Among them, $Y_{\min}$ belongs to the set minimum value, and usually its value range is $g\,(a) + Y_{\min} > 0$.

$$\text{Fitness}\,(g\,(a)) = \begin{cases} Y_{\max} - g\,(a), \\ 0. \end{cases}\qquad(16)$$

Among them, $Y_{\max}$ belongs to the set maximum value, and usually its value range is $Y_{\max} - g\,(a) > 0$.

$$\text{Fitness}\,(g\,(a)) = \frac{1}{1 + k - g\,(a)}.\qquad(17)$$

Formula (17) represents the maximum optimization problem function expression, where $k - g\,(a) \ge 0$.

$$\text{Fitness}\left(g(a)\right) = \frac{1}{1 + k + g(a)}. \tag{18}$$

Formula (18) represents the minimum optimization problem function expression, where $k + g(a) \geq 0$.

According to the algorithm of the objective optimization problem, we can use it in the vocal music teaching classroom. We regard the different characteristics of students as different optimization objectives. Using multi-objective optimization, we can find a teaching algorithm suitable for all students, so that the results of acoustic teaching can be achieved.

3.2. Vocal Music Teaching. As an ancient art, vocal music is a combination of music and language. Since the human voice is added to the process of vocal music performance, the performance effect of vocal music is more moving [14]. Vocal music teaching practice activities are social, and teaching methods are unique, cooperative, and practical. Articulation is the most important thing during a vocal performance. It directly determines the effect of the performance to some extent. In related works, it has been described as "the clearness of the word is one must; the purity of the cavity is the second; the board is the three". "Banzheng" refers to the need of pronouncing clearly. It should first be the correct character, so that the audience can clearly hear what is being sung. This also shows the importance of language in vocal music. However, according to the current market situation, it is not difficult to find that it is still a problem that many singers need to face. This problem is mainly caused by the lack of basic vocal skills, the lack of research on the standardization of Chinese pinyin, and the failure to connect the sound of the word and the melody line of the music. In response to these problems, relevant scholars believe that they can be solved from the aspects of phoneme and word orthography. Figure 3 shows the structure of audio data processing.

Breathing is the foundation and motivation of singing and vocalization. From a professional level, the breath directly affects the performance of the performance. Exhalation is the expulsion of air out of the body. Inhalation means taking air into the body. These two actions are opposite and happen in succession at the same time. If there is excessive breathing during singing, it will reduce the integrity of the work. In order to improve the level of the work, it is necessary to train the breathing method when singing. In general, it is necessary to consciously control the breathing rate according to the length of the phrase to prevent the organ from becoming rigid [15]. Figure 4 shows a training diagram of the vocal pronunciation system structure.

If vocal skills are directly transmitted in language, it will be very obscure and difficult to understand. If it can be combined with visual thinking when explaining, it will make the skills more vivid, arouse students' interest, and make students master more accurately. Therefore, combining teaching with visual thinking is one of the most effective teaching methods.

3.3. Visual Thinking. Thinking refers to thought or thinking, which is a high-level physiological phenomenon. It belongs to the biochemical reaction in the brain and is a high-level stage of human cognitive process. In the process of vocal music singing, it is also necessary to carry out thinking activities, and it is necessary to use thinking to lead one's own consciousness and complete the work. In vocal singing, there are action thinking, image thinking, convergent thinking, and divergent thinking. In the practice of singing training, action thinking and image thinking play an important auxiliary role in the improvement of our vocal music practice. For some specific song singing behavior, the singer needs to analyze it with the help of thinking. For example, we need to use hearing to judge the beauty of sound. This is the most important part of thinking. For beginners in vocal music, it is necessary to strengthen them. This makes it gradually form a correct sound image thinking, which lays a good foundation for future learning. Figure 5 shows the simple vocal practice structure.

Visualization is a means of processing images and data using computer means. It is a way to display information more intuitively. The technology was first used in the computer field. In terms of how it works, visualization is the process of turning information into images. Thinking visualization is the reflection of invisible ways of thinking through images. In fact, we often use its methods in daily life. Thinking visualization refers to presenting thinking methods and thinking paths that were otherwise invisible using a series of graphic techniques to make them clearly visible. In fact, we often use its methods in our daily life. For example, we often say we can do it by bending the calf, but we do not know which way to bend it. At this time, if someone makes a demonstration, it can be displayed intuitively. This is a process of thinking visualization. This method is also very common in vocal music teaching. Figure 6 shows the basic structure diagram of visualization technology.

4. Application Experiment

Vocal Learning Data Collection: in order to investigate the current state of vocal music learning, we first need to conduct a critical survey of student learning. In this experiment, we used a questionnaire survey to investigate the vocal music students in place A. The details are as follows:

In order to analyze the students' situation, we have investigated the students' comprehensive quality of vocal music. According to the actual situation of students, their comprehensive quality is divided into four grades. According to the data in Table 1, there are 37 students with good overall quality, accounting for 27.6% of the total number of respondents. There are 31 students with average comprehensive quality, accounting for 23.1% of the total number of respondents. There are 46 students who have passed the comprehensive quality, accounting for 34.3% of the total number of respondents. There are 20 students with unqualified comprehensive quality, accounting for 15% of the total number of respondents. According to the survey data, there is a big gap in the comprehensive quality of students. Generally speaking, the level is relatively average,
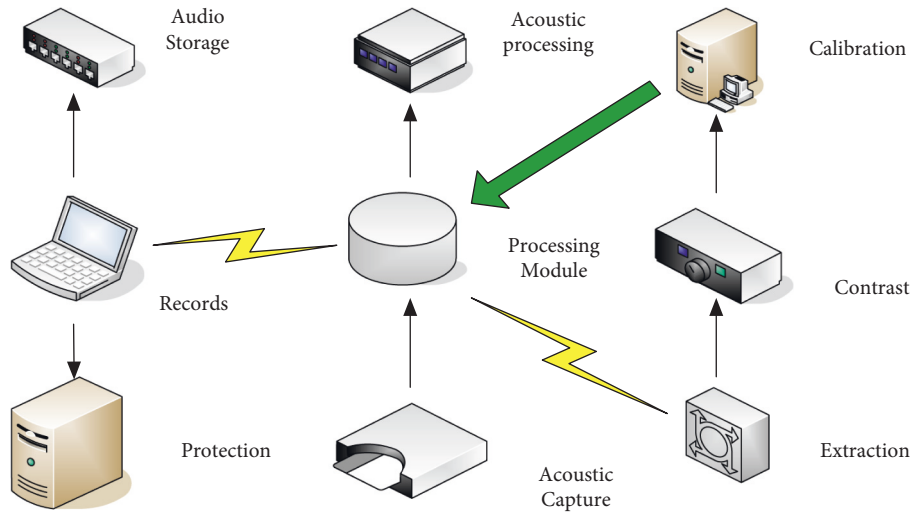
FIGURE 3: Audio data processing structure.
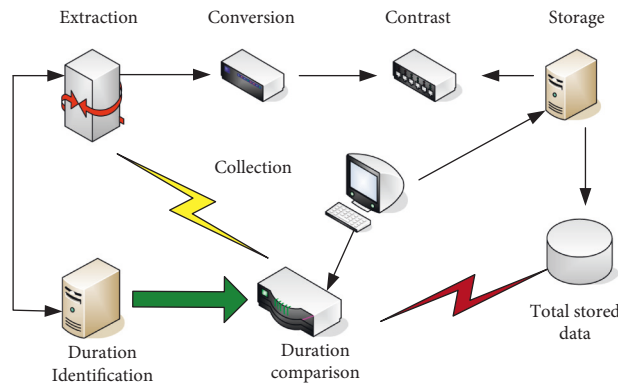
FIGURE 4: Vocal articulation system structure training.
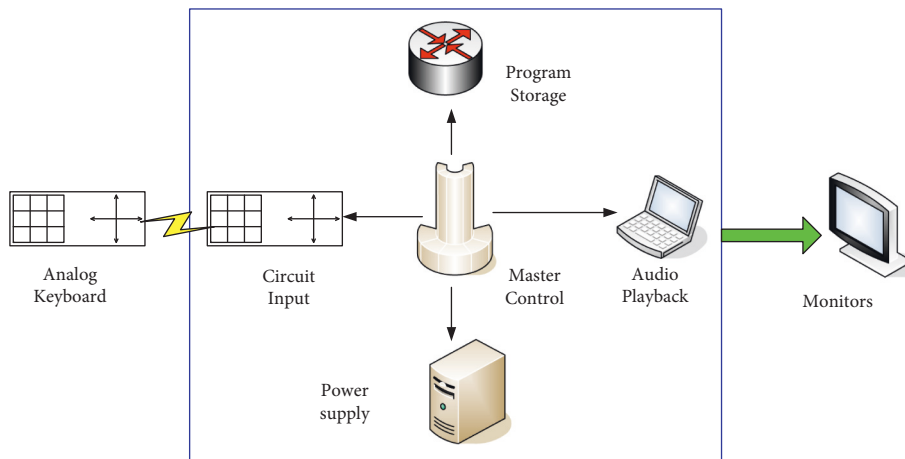
FIGURE 5: Simple vocal practice structure.

the students' cultural background is not very high, and the cultural background and accomplishment are relatively weak.

According to the data in Table 2, we have investigated and analyzed the students' vocal music foundation and mastery of skills. According to the data, there are 6 people

with the ability to teach acoustics, accounting for 4.5% of the total number of respondents. There are 23 people with singing experience, accounting for 17.2% of the total number of respondents. There are 39 people who can master the knowledge and skills proficiently, accounting for 29.1% of the total surveyed. 66 people have basic knowledge and skills,
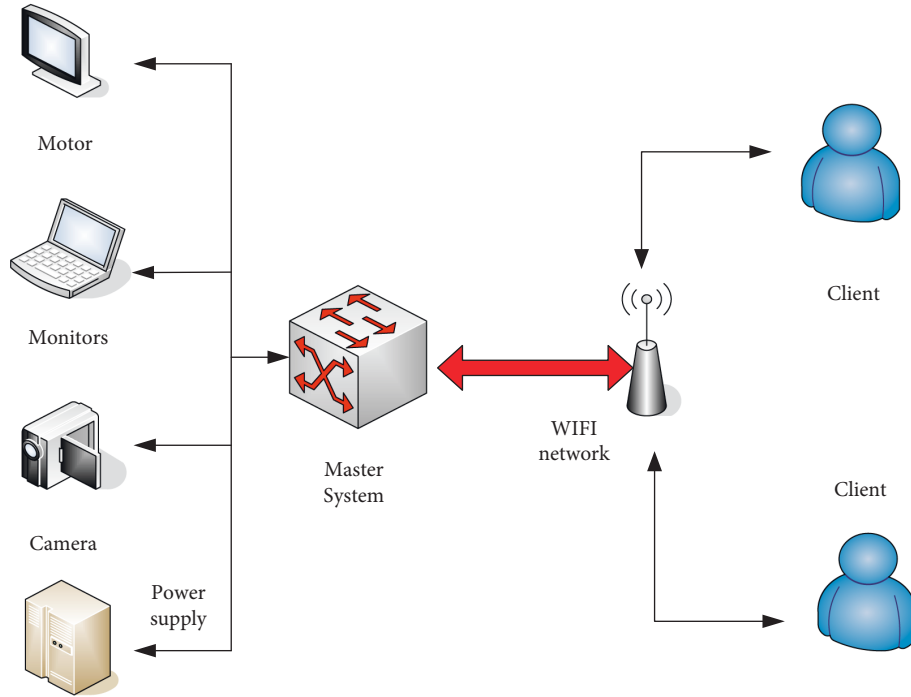
FIGURE 6: Basic structure of the visualization technique.

TABLE 1: Analysis of students' comprehensive quality survey.

Category	Number of people	Proportion
Good	37	27.6
Moderate	31	23.1
Pass	46	34.3
Not passing	20	15

accounting for 49.3% of the total surveyed. According to the data, the proportion of students with basic mastery is the highest, and the proportion of students with vocal music teaching ability is the lowest. It can be seen from this that students generally lack practical experience and have fewer opportunities for vocal music activities. According to the actual situation, most of the students pay attention to the learning skills, lack the perception of the works, and have a low degree of completion.

Experimental Data: the theme of this study is to improve the effect of vocal music teaching with the help of a genetic algorithm. In order to investigate the effect of genetic algorithm, we conduct a brief data analysis of the results of different calculation methods of genetic algorithm. The details are as follows:

From the experimental data in Table 3, we can see that although the algorithm NSGA-II is stronger than the algorithm USMOEA in terms of breadth, the original intention of the experimental design is to obtain a nondominated solution set that is closer to the Pareto front and more uniformly distributed. Then from the data in the table, the expected solution goal has been basically achieved. The algorithm can obtain a nondominated solution.

From the experimental data in Table 4, we can see that the values of ZDT1, ZDT2, and ZDT3 are compared.

TABLE 2: Analysis of the mastery of basic vocal knowledge and basic skills.

Category	Number of people	Proportion
Ability to teach vocal music	6	4.5
Experience in singing	23	17.2
Proficiency in	39	29.1
Basic mastery	66	49.3

TABLE 3: Algorithm NSGA-II and algorithm USMOEA on the M-measure of the test function.

M-measure	Group	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
	BEST	1.2984	1.0320	1.7543	1.2198	0.7816
USMOEA	MEAN	1.0569	0.0599	1.5328	0.6692	0.7569
	WORST	0.8149	0.8174	1.0542	0.5473	0.6736
	BEST	1.4159	1.3298	2.3029	1.3036	0.8851
NSGA-II	MEAN	1.2578	1.0798	1.7319	0.8132	0.8269
	WORST	0.9899	0.9638	1.2978	0.7056	0.7025

TABLE 4: Algorithm NSGA-II and algorithm UCMOEA on the S-measure of the test function.

S-measure	Group	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
	BEST	0.0027	0.0021	0.0099	0.0163	0.0011
UCMOEA	MEAN	0.0214	0.0031	0.0196	0.0196	0.0039
	WORST	0.0299	0.0135	0.0213	0.0235	0.0047
	BEST	0.0036	0.0040	0.0100	0.0048	0.0016
NSGA-II	MEAN	0.0319	0.0075	0.0236	0.0089	0.0035
	WORST	0.0736	0.0120	0.0639	0.0132	0.0048

Algorithm NSGA-II is weaker than algorithm UCMOEA in terms of worst data, mean, or best data. According to the data obtained by ZDT4, we can see that the algorithm NSGA-II is stronger, and the algorithm UCMOEA is relatively weak. For the experimental data of ZDT6, we can see that the algorithm NSGA-II is better than the algorithm UCMOEA in the mean and worst cases, but the algorithm UCMOEA is stronger than the algorithm NSGA-II in the optimal data.

Attitudes of Students toward Vocal Music Courses: with the continuous progress of national education, the education curriculum has become more and more perfect, and the requirements for comprehensive quality have become more and more strict. In addition to the strict requirements for subject education, we also pay more and more attention to art courses, incorporating the comprehensive development of students into the educational requirements. In order to explore students' attitudes toward vocal music courses, we conducted a brief analysis of students' attitudes toward vocal music courses in place A. The details are as follows:

To explore why students take vocal lessons, we have grouped them into four broad categories. According to the data in Figure 7, there are 22 people who choose vocal music courses purely for learning needs, accounting for 16% of the total number of respondents. 19 people chose vocal music courses because of their personal hobbies, accounting for 14% of the total surveyed. There are 31 people who choose vocal music courses to pave the way for later work, accounting for 23% of the total surveyed. There are 63 people who choose vocal music courses to learn music knowledge and vocal skills, accounting for 47% of the total surveyed. According to this data, most students choose vocal music classes to learn vocal skills. This shows the importance of vocal skills in vocal music courses.

To understand the school's vocal learning climate, we surveyed students' personal perceptions. According to the survey data, 76 people believe that the atmosphere for learning vocal music in schools is very good, accounting for 57% of the total number of respondents. 19 people think that the atmosphere of learning vocal music in schools is average, accounting for 36% of the total number of respondents. 14 people think that the atmosphere for learning vocal music in schools is relatively poor, accounting for 11% of the total number of respondents. There are also 6 people who do not know about the school's vocal music atmosphere, accounting for 5% of the total number of respondents. According to the survey data, most students believe that the learning atmosphere is good, indicating that students are generally more serious about the course.

In order to improve the teaching effect, we use different genetic algorithms to optimize the design. As can be seen from the data in Figure 8, from the perspective of breadth, the effect of the algorithm UCMOEA is not particularly good. But the original intention of the experiment is to obtain a nondominated solution set that is close to the Pareto front and has a more uniform distribution. The data in Table 4 show that the expected solution goal has been

achieved. Algorithm UCMOEA can obtain the nondominated solution set with a more uniform distribution and closer to the Pareto front.

By testing several functions with different properties, the results show that both the algorithm UCMOEA and the algorithm USMOEA have good search ability compared with NSGA-II. The set of nondominated solutions obtained by the algorithm designed in this study is closer to the Pareto boundary, and the distribution of nondominated solutions is more uniform. It can be seen that in the actual vocal music teaching process, teachers can adopt this algorithm to improve the teaching effect.

Survey and Analysis of Vocal Music Course Content: reasonable arrangement of the course content will help students to concentrate in class, improve classroom efficiency, and master theoretical knowledge and skills more accurately. In order to understand the arrangement of vocal music courses in place A, we conducted a random survey of students. The investigation is as follows:

According to the arrangement of the course, we divide it into four levels. According to the data in Figure 9, there are 27 people who think the course arrangement is very reasonable, accounting for 20% of the total number of respondents. There were 31 people who thought the course arrangement was more reasonable, accounting for 23% of the total number of respondents. There were 19 people who thought that the course arrangement was very general, which accounted for 14% of the total number of respondents. There were 58 people who thought the course arrangement was very unreasonable, which accounted for 43% of the total surveyed. According to the data, the number of people who do not agree with the curriculum arrangement is the highest. It can be seen that students generally believe that the school's curriculum needs to be adjusted.

In addition to the classroom arrangement, the teaching content of the vocal music classroom is also very important. In order to understand the content of local vocal music teaching, we conducted a brief survey of it. According to the survey data, there are 45 people who think that songs should be appreciated in the classroom, accounting for 34% of the total number of respondents. There were 28 people who thought that song training should be carried out in the classroom, accounting for 21% of the total number of respondents. There were 52 people who thought that the classroom should be tested, accounting for 39% of the total number of respondents. Eight other students indicated that other methods could be taken. From these data, it can be seen that most people generally hope that they can carry out actual testing and training in the classroom to correct their deficiencies in time.

Investigation and Analysis of Classroom Teaching Content Types: vocal music classes need to appreciate and train music in the classroom. To understand student attitudes toward classroom instruction, we surveyed the types of music students would like to see in the classroom. The details are as follows:

According to the data in Figure 10, students want teachers to choose a variety of different types of music mixed teaching in the classroom, accounting for 76% of the total
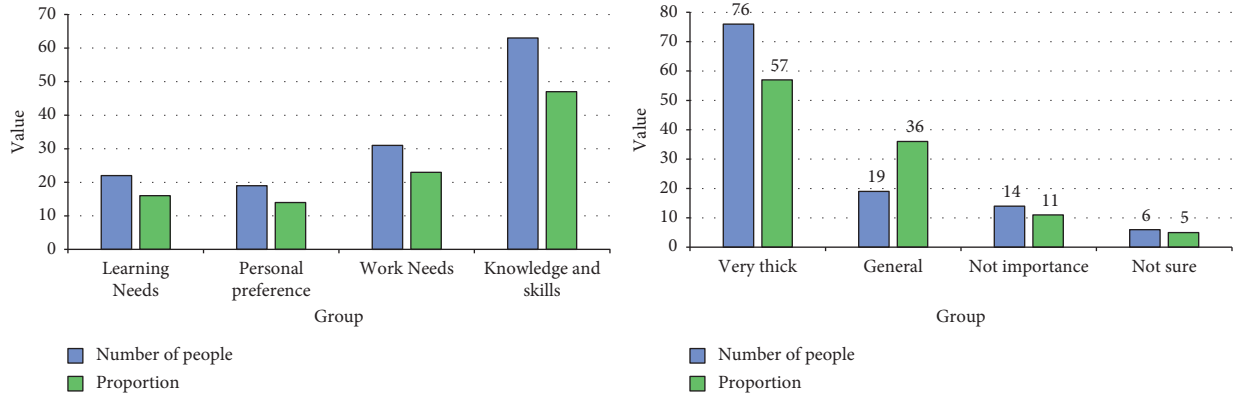
FIGURE 7: Analysis of students' attitudes toward vocal courses.
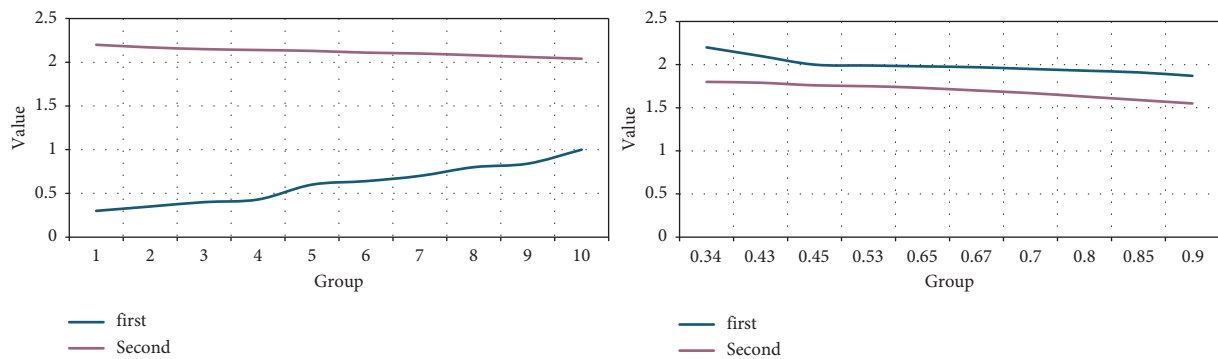
FIGURE 8: Comparison of the results of the two algorithms for ZDT1 after one run.
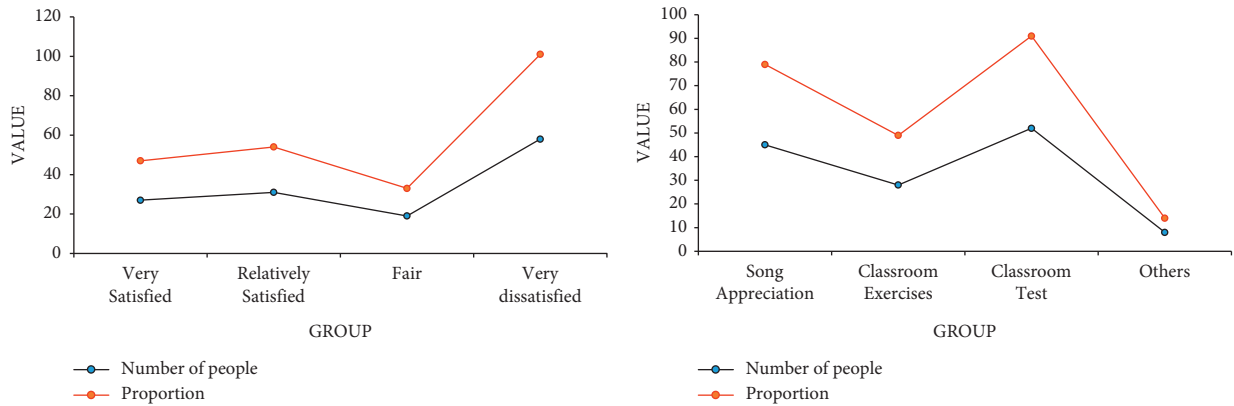
FIGURE 9: Analysis of vocal course content survey.

number of respondents. 16 students indicated that they can use children's repertoire properly in the classroom, accounting for 12% of the total surveyed. 9 students hope that teachers can choose popular songs as training repertoires in the classroom, accounting for 7% of the total number of respondents. 6 students hope that teachers can choose popular and folk songs as training repertoires in the classroom, accounting for 5% of the total number of respondents. According to these data, most students hope that the training repertoire in the classroom can be diversified.

Vocal practice is the basic learning of vocal music. When we surveyed vocal practice, we found that 94 students thought vocal practice was very necessary, accounting for 70% of the total surveyed. Twenty-five students considered vocal practice necessary, accounting for 19% of the total surveyed. Twelve students considered vocal practice to be optional. Three other students felt that vocal training was unnecessary. According to the data, most students attach great importance to vocal practice. However, there are still some students with cognitive biases that need to be corrected.
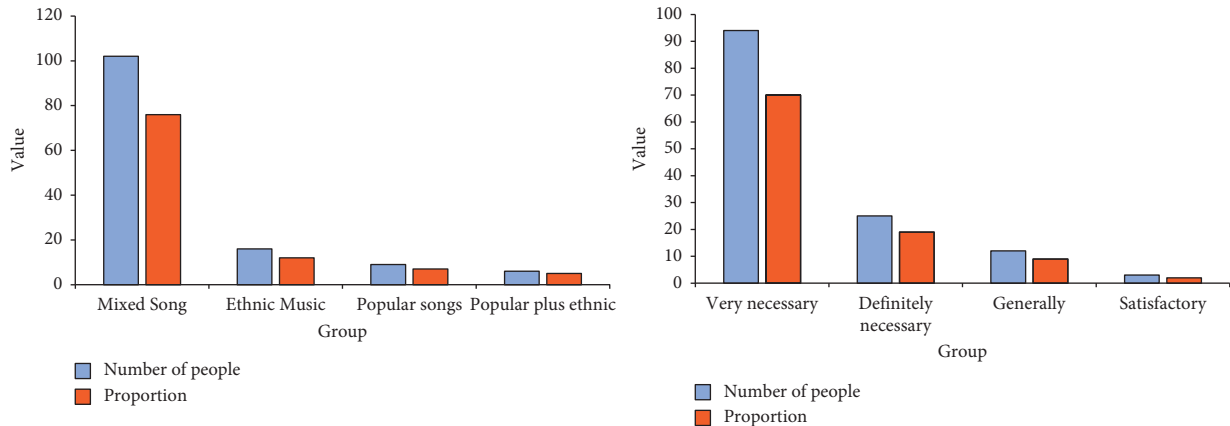
FIGURE 10: Analysis of classroom teaching content type survey.

5. Conclusion

This study discusses the current deficiencies in the field of vocal music education and uses genetic algorithms and visual thinking technology to combine vocal music teaching, so that the teaching methods are in line with the characteristics of students and improve teaching efficiency. Vocal music teaching is an important part of music art, and improving vocal skills is of great help to music singing. In the existing market, although various singing methods emerge in an endless stream, there are still many problems in the singing of songs. For example, unsteady breathing. This study points out that, in the process of vocal music teaching, attention should be paid to the teaching atmosphere, teachers' teaching methods, and teaching content, especially the students' attitude toward vocal music teaching, so as to correct the problem ideologically. By combining genetic algorithm and visual thinking with vocal music teaching, the existing teaching method is improved, so that the teaching method is consistent with the actual situation, which is conducive to the intuitive display of vocal music skills. This method provides a reference for future vocal music teaching methods, which is conducive to improving teaching results and improving vocal music skills. Although this study discusses vocal music teaching to a certain extent, it still has shortcomings: due to the lack of personal experience, many studies are only superficial research, just a simple overview of singing skills.

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.

References

[1] R. Tavakkoli-Moghaddam, J. Safari, and F. Sassani, "Reliability optimization of series-parallel systems with a choice of redundancy strategies using a genetic algorithm," *Reliability Engineering & System Safety*, vol. 93, no. 4, 2017.

[2] H. Dawid and M Kopel, "on economic applications of the genetic algorithm: On economic applications of the genetic algorithm: a model of the cobweb type," *Journal of Evolutionary Economics*, vol. 8, no. 3, pp. 297–315, 1998.

[3] A. Volkanovski and B. Mavko, "Genetic algorithm optimisation of the maintenance scheduling of generating units in a power system," *Reliability Engineering & System Safety*, vol. 93, no. 6, pp. 779–789, 2017.

[4] Q. Long and C. Wu, "A hybrid method combining genetic algorithm and Hooke-Jeeves method for constrained global optimization," *Journal of Industrial and Management Optimization*, vol. 10, no. 4, pp. 1279–1296, 2014.

[5] Y. Yoshitomi, H. Ikenoue, T. Takeba, and S. Tomita, "GENETIC algorithm in uncertain environments for solving stochastic programming problem," *Journal of the Operations Research Society of Japan*, vol. 43, no. 2, pp. 266–290, 2000.

[6] M. H. Alavidoost, M. H. F. Zarandi, M. Tarimoradi, and Y. Nemati, "Modified genetic algorithm for simple straight and U-shaped assembly line balancing with fuzzy processing times," *Journal of Intelligent Manufacturing*, vol. 28, no. 2, pp. 313–336, 2017.

[7] N. Metawa, M. K. Hassan, and M Elhoseny, "Genetic algorithm based model for optimizing bank lending decisions," *Expert Systems with Applications*, vol. 80, no. 9, pp. 75–82, 2017.

[8] D. W. Gong, J. Sun, and Z Miao, "A set-based genetic algorithm for interval many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 47–60, 2018.

[9] T. Fetouh and M. S Zaky, "New approach to design SVC-based stabiliser using genetic algorithm and rough set theory," *IET Generation, Transmission & Distribution*, vol. 11, no. 2, pp. 372–382, 2017.

[10] K. Rajeswari and S Neduncheliyan, "Genetic algorithm based fault tolerant clustering in wireless sensor network," *IET Communications*, vol. 11, no. 12, pp. 1927–1932, 2017.

[11] S. Prichard, "A mixed-methods investigation of preservice music teaching efficacy beliefs and commitment to music teaching," *Journal of Research in Music Education*, vol. 65, no. 2, pp. 237–257, 2017.

[12] N. Steele Royston and Natalie, "Improving music teaching through interpersonal relationships," *Music Educators Journal*, vol. 103, no. 4, pp. 34–39, 2017.

[13] A. Gholam, "Visual thinking routines: classroom snapshots," *Athens Journal of Education*, vol. 6, no. 1, pp. 53–76, 2019.

[14] D. He and N. Luo, "Retraction Note: Spatiotemporal evolution characteristics of extreme rainfall based on intelligent recognition and evaluation of music teaching effect in colleges and universities," *Arabian Journal of Geosciences*, vol. 14, no. 23, p. 2491, 2021.

[15] A. Remacle, M. Garnier, S. Gerber, C. David, and C Petillon, "Vocal change patterns during a teaching day: inter- and intra-subject variability," *Journal of Voice*, vol. 32, no. 1, pp. 57–63, 2018.

Hindawi

Research Article

Improved Chaotic Algorithm-Based Optimization Technology of Architectural Engineering Drawing Parameters

Xiaoqiu Ma (ID)

Department of Basic Science, Jilin Jianzhu University, Changchun 130000, China

Correspondence should be addressed to Xiaoqiu Ma; maxiaoqiu@jlju.edu.cn

The traditional methods deal with large sample data sets of architectural engineering drawings and they have high time complexity and space complexity as well. Their searching time is long and sometimes the results are unsatisfactory. Therefore, this paper proposes an optimization method designed for architectural engineering drawing parameters to overcome the limitations of the traditional methods. It is based on the improved chaotic algorithm. The algorithm proposes the optimization model of architectural engineering drawing (AED) parameters in the first phase. In the second phase, an improved chaos algorithm is used to optimize the parameters of architectural engineering drawing, and the modeling strategy of visual parameter optimization environment is constructed. Finally, the visualization parameter optimization process of architectural engineering drawing is completed. Through experiments, it is evidently observed that the method presented in this paper can effectively reduce the optimization time, improve the lighting illumination of buildings, and improve the optimization precision of architectural engineering drawing parameters. The proposed method considers multiple parameters and it has greater application ability in the field of architectural designs.

1. Introduction

1.1. Background Study. In recent 20 years, the parametric design of building engineering has gradually become a new topic in the field of design. Parametric design transforms the concern of architectural design result into the concern of design process and control. Parametric design gets rid of the traditional modeling methods of sketch initiation or pattern reproduction, but controls the establishment of design schemes through the modeling and optimization of parametric models [1]. The process of parametric design can be divided into site investigation, form-finding, design evolution, structural system, and structural logic [2]. The parametric design process and scheme generation are characterized by complexity and nonlinearity. Parametric design focuses on the information carried by architectural elements and follows certain norms and rules to create architecture from the bottom up [3]. Parametric design techniques impart advantages for architecture design processes. Architects apply these methods in their creation

of design suggesting solutions at an earlier stage of the process to speed up the design process and enhance the accuracy [4].

1.2. Existing Literature. In [4], the authors propose a process based on parametric design creation. It combines architectural design with parametric modeling methods. This makes the design objectives more comprehensive and also supports the designers in finding architectural solutions. In [5], the authors propose an optimization algorithm for architectural engineering drawing parameters based on Spark parallel SVM. The Spark cluster was used to broadcast the training set of architectural engineering drawing parameters to each Executor in the form of variables. The parameter optimization process of parallel SVM is carried out. Task parallelism is applied in order to speed up the optimization speed. In [6], the researchers put forward building cover day feed parameters based on the firefly algorithm of intelligent optimization method. It improves the HATA experience

model covering the simulation of the building. It shows good processing precision. In [7], the authors propose an algorithm study based on the adaptive determination of DBSCAN algorithm parameters. It generates candidate EPS and other parameters using the distribution characteristics of the data set itself. The algorithm can realize the full automation of clustering process and can choose reasonable EPS and other parameters. It shows high clustering accuracy. In [8], the authors state that the parametric optimization of an architectural object is an important design task. They also state that improvement is needed in the area of quality of a design solution. Also, there is a need to take care of energy efficiency in design. The authors compare the calculated compactness ratios of the external envelope for different storeyed buildings. In [9], the authors describe different approaches for the parameterization in architecture. They also provide the comparison of these approaches. They map the design process to parametric data flow in each of the approaches. In [10], the authors propose two approaches for parametric design creation (1) Creating models in architectural design and (2) Using various tools of applied mathematics. They state that both the approaches need optimization for the purpose of modeling in architectural design. Authors use an interdisciplinary approach for identifying the role of optimization in architectural design.

1.3. Contributions of the Paper. From the existing literature, it can be concluded that though there is a lot of research in the field of Parameters Optimization Model for Architectural Engineering Drawing, they have the problem of poor prediction accuracy. Some of the methods show good accuracy but have high processing time. This paper proposes an optimization technique of architectural engineering drawing parameters based on an improved chaotic algorithm. Our aim is to design a Parameter Optimization Model for Architectural Engineering Drawing which is accurate and faster as compared to the previous methods.

1.4. The Major Highlights Are as Follows:

(1) Proposing parametric optimization model which is novel in idea.

(2) Visual parameter optimization condition design is presented.

(3) Parameter optimization process is devised.

(4) Proposing the optimization method of building engineering drawing parameters using an improved chaos algorithm.

The following sections explain the proposed work.

2. Construction of Parameters Optimization Model for Architectural Engineering Drawing

This is a three phase process in which, first of all, find a set of optimal parameter values. In the next phase, that is, visual parameter optimization condition design, parameters that satisfy the specific conditions are chosen. These conditions

have been mentioned in detail in Section 2.2. In the final phase, Parameter optimization process is explained.

2.1. Parametric Optimization Model. Parametric optimization is to find a set of optimal parameter values $\alpha* \in \alpha$ in the feasible region of the design vector $\alpha = [T_1, T_2, T_3, ..., T_{n-1}, T_n]^T$, so that the objective function $J = f(\alpha*)$ takes the minimum (or maximum) value. For the parameter optimization problem [7], if the design variable α is a set of parameters, then the design problem in r dimensional space can be discussed as in equation (1).

$$\alpha\prime = [T_1, T_2, T_3, ..., T_r]^T. \tag{1}$$

It is clear that at this point α' is a fixed point in the Euclidean vector space E^r (design space) composed of α, which can be expressed as $\alpha \in E^r$. When the design variable is a parameter, the objective function can be described as in .

$$J = f(\alpha) = f(T_1, T_2, T_3, ..., T_r). \tag{2}$$

According to equations (1) and (2), the general form of the parametric optimization mathematical model can be expressed by equation (3).

$$\min_{\alpha \in D} f(\alpha), \tag{3}$$

where the objective function $f(\alpha)$ is a real-valued function, and D is the domain of $f(\alpha)$. Equation (3) can also be written in a more general form as the following:

$$\begin{cases} \min_{\alpha \in D} f(\alpha)' \\ s.t.\, d(\alpha) = 0 \\ h(\alpha) \le 0, \end{cases} \tag{4}$$

where $\alpha' = [T_1, T_2, T_3, ..., T_r]^T$ is an r dimensional design vector, $d(\alpha) = [d_1(\alpha), d_2(\alpha), ..., d_p(\alpha)]^T$ is a p dimensional equality constraint, and $h(\alpha) = [h_1(\alpha), h_2(\alpha), ..., h_q(\alpha)]^T$ is a q dimensional inequality constraint.

In equation (3), since the design variable is unconstrained to take all the values in the domain D, it belongs to the unconstrained extremum problem. In equation (4), $\alpha \in D$ is constrained by equations (5) and (6), that is,

$$d_i(\alpha) = 0, i = 1, 2, .., p, \tag{5}$$

$$h_j(\alpha) \le 0, j = 1, 2, .., q. \tag{6}$$

Therefore, equation (4) is the constraint extremum condition.

2.2. Visual Parameter Optimization Condition Design. For the optimization of visual parameters, the following conditions should be defined:

(a) One of the key problems in optimizing the visual parameters of the mathematical model and graphical transformation of the system is in what form the mathematical model can be graphically or visually transformed [11]. In fact, the equation constraint of

optimization problems, such as differential equation, difference equation, transfer function, equation of state, and other forms, is visualized by graphical display technology according to the modular modeling idea.

(b) The feasible region is the set of points in the design space that satisfy all the constraints.

(c) Excitation function of the system includes instruction function and interference function. In order to facilitate the comparison of dynamic characteristics of the system, some typical signal functions are generally selected as instruction signals, such as step function, impulse function, and random signal [12, 13].

(d) For example, the initial time t_0 of the end and terminal conditions, the terminal time t_f, the initial state $x(t_0)$, and the terminal state $x(t_f)$ are either given or arbitrarily valued according to the task conditions.

(e) For visual parameter optimization problems, the objective function is generally shown in equation (2).

2.3. Parameter Optimization Process. After the designers analyze and calculate a variety of architectural form design schemes, the final form scheme is obtained. From the perspective of natural lighting and ventilation, the position and size of the internal atrium of the building are further optimized [14]. The actual parameter optimization problem is usually solved by numerical iteration method [15].

For general (multi-objective) optimization problems, the algorithm flow is as follows:

In the first step, initial design vector α_0 was selected. In the second step, feasible direction s_0 was found. In the third step, step size factor l_0 is selected along the feasible direction, and the new design vector is obtained from s_0 and l_0 as the following equation (7).

$$\alpha_1 = \alpha_0 + l_0 s_0. \tag{7}$$

The fourth step is to find a new design direction s_1 and step length l_1 from α_1. Iterate one step forward from point α_1 to point α_2, and so on, many iterations. The iteration form from step k to step $(k+1)$ is shown in equation (8).

$$\alpha_{k+1} = \alpha_k + l_k \mathbf{S}, \tag{8}$$

$$f(\alpha_{k+1}) < = f(\alpha_k), k = 0, 1, 2, \dots. \tag{9}$$

The fifth step is to check whether the preset accuracy is reached at each iteration step. If it is, the minimum point is considered to have been found; otherwise, the iterative calculation should be continued. In this paper, equation (9) is taken as the criterion for the termination of iteration as shown in equation (10)

$$\frac{\|\alpha_{k+1} - \alpha_k\|}{\|\alpha_k\|} \le X, \tag{10}$$

where X is the given precision.

3. Optimization Technology of Building Engineering Drawing Parameters Based on Improved Chaotic Optimization Algorithm

The parameter optimization technique is applied for engineering drawing parameters. The steps are as follows:

(1) Proposing an Optimization method of building engineering drawing parameters

(2) Identifying the modeling strategy of visual parameter optimization environment.

(3) Proposing the visualization of parameter optimization process, that is, identifying the optimal parameters.

(4) Applying the chaos algorithm for verifying the optimization effect of this method on the parameters of architectural engineering drawing

(5) Comparing the results with existing algorithms like SPARK SVM, firefly algorithm, and DBSCAN algorithm

In the design of visual parameter optimization environment, the following is chosen.

3.1. Optimization Method of Building Engineering Drawing Parameters. Chaos is a common phenomenon existing in nonlinear systems, and its motion is characterized by ergodic, randomness, and regularity [15]. Logistic mapping is a successful example of chaos research in nonlinear equations. It was originally used to describe the generation change rule of insect numbers. Logistic mapping is shown in equation (11).

$$x_{n+1} = \mu x_n (1 - x_n) n = 0, 1, 2 \dots, \tag{11}$$

where μ is the control parameter, and n is a normal number. When $3.544090 \le \mu \le 4$, a chaotic trajectory can be iterated from any initial value $x_0 \in (0, 1)$, which is traversed within the range of (0,1). The chaos optimization algorithm takes advantage of this traversal feature, and the basic steps are as follows:

One, in formulas (1) and (13), vector x_n is, respectively, endowed with m initial values with slight differences (not 0.25, 0.5, and 0.75), and m is the number of parameters to be optimized. After k iterations, m different chaotic variables $x_i(k) i = 1, 2, \dots, m$ can be obtained.

Two, by mapping the chaotic variable $x_i(k)$ to the search space of the optimization variable, $X_i(k)$ can be obtained:

$$X_i(k) = c_i + d_i x_i(k), \tag{12}$$

where, c_i and d_i are constants related to the variable search space $[a, b]$, $c_i = a, d_i = b - a$. For the first chaotic optimization search (rough search), the optimization index is $\min f(X_i)$. Let $f *$ be the current optimal solution, and $X *$ be the corresponding parameter combination. $f(X_i(k)) < f *$, Substitute $X_i(k)$ into the calculation optimization index $f(X_i(k))$ successively, $f(X_i(k)) < f *$, then equation (13) is obtained.

$$f* = f(X_i(k)), X* = X_i(k). \tag{13}$$

Otherwise, $X_i(k)$ is discarded. Until $f*$ star stays the same for some number of steps.

For the second chaotic optimization search (fine search), according to equation (14), a new chaotic variable $X_i(k')$ is generated, and a_i is the control variable less than 1.

$$X_i(k') = X_i^* + \alpha_i(x_i(k') - 0.5). \tag{14}$$

Three, put $X_i(k')$ in order to calculate the optimization index $f(X_i(k'))$. If $f(X_i(k')) < f*$, then $X* = X_i(k')$, $f* = f(X_i(k'))$. Otherwise, give up $X_i(k')$. Until $f*$ remains unchanged for several steps, the result is output.

3.2. Modeling Strategy of Visual Parameter Optimization Environment. In the design of visual parameter optimization environments, MATLAB 5.2 is utilized as the numerical calculation engine of optimization operation, and uses SIMuLink K 2.2 as the visual modeling support environment of nonlinear parameter optimization systems. The modeling idea is based on the separability of the system, that is, the large-scale system can be divided into several subsystems, and the subsystems can be decomposed into more primitive subsystems. Subsystem itself is defined as a visual component or module with independent operation function, which generally includes several input and output ports and is connected with each other through data flow channel; its function is to transform the incoming data or signal. The way to construct the visual optimization model is to establish the overall optimization simulation model by connecting the component models (submodels or modules) that constitute the system model. If the component model itself is constructed from a more primitive component model, it can form a hierarchical or hierarchical structure of the visual model. Therefore, the visual module (component) graphic modeling system based on SIMULINK is an important part of the advanced visual parameter optimization environment. It can provide users with a flexible, fast, and easy to expand high-level visual integrated information processing platform [16].

One, the graphical modeling components of SIMU-LINK are directly connected with each other, and the interfaces with the submodules of each layer of the model are retained in the visual environment.

Two, for the complex nonlinear parameter optimization model, the separation system technology is used to decompose the complex model into several subsystems at the next level according to the function. For the large subsystem, it can be further decomposed into subsystems at the next level. For this reason, the top-down hierarchical structure of simulation is established by analogy.

Three, it enhances the ability of human-computer interaction in the process of optimization, enables the user to pause or terminate the iteration at any time in the process of optimization, and reprocesses the data after optimization.

Four, it allows users to modify the hierarchical structure, parameters, and iteration initial values of the optimization model at any time, and allows users to run and test the subsystems of each layer, respectively, which is convenient for users to optimize and locate the errors of the subsystems of each layer.

Five, subsystem reuse technology allows users to design (subsystem) or document can be reused many times.

Six, users can customize, modify, or expand the optimization model library and algorithm library. The six considerations can ensure that the optimization system has clear structure, is easy to modify independently, is easy to debug and maintain, and has good scalability. The visual optimization environment is mainly composed of the following six parts through seamless links:

(a) The model library provides a visual modeling platform and auxiliary graphic modeling tools through graphical mode. The model component is composed of model library function modules in SIMuLinK 2.2, which is used to complete the connection, conversion, calculation, and result output of system simulation optimization model.

(b) By calling MATLAB optimization toolbox, the optimization algorithm library can provide users with a variety of optimization algorithms for solving such problems as parameter optimization, unconstrained optimization, quasi Newton realization, least square optimization, nonlinear least square realization, constrained optimization, SQP realization, and multi-objective optimization. The optimization algorithm can also be designed and developed by users themselves.

(c) Model management optimization model management includes simulation layer management and application layer management. Simulation layer management is used to create, modify, open and access system models, and define and modify submodel blocks, including parameters, types, and attributes; application layer management is used to manage the files of the models according to the optimization tasks, M-file and MDL-file are used to share resources between models or platforms

(d) Optimization database management takes MAT-LAB language as host language to help users manage, access, retrieve, and analyze the data in the optimization process. It can also dynamically analyze the optimization process, dynamically modify model parameters with optimization data, and compare the results of system optimization under different conditions.

(e) The optimization environment settings are mainly used to set foreground, foreground and relation types (line type, line width, color, etc.) in the visual modeling area, as well as the running environment settings

(f) The virtual oscilloscope technology is applied in real-time monitoring, which can monitor and analyze the waveform of any number of ports (input or output) of the optimization model at any stage and any moment of the optimization process.

3.3. Visualization of Parameter Optimization Process. As mentioned in Section 3.2, the modeling and simulation strategy of SIMuLinK can be adopted to establish the visual parameter optimization environment of general nonlinear system, and the basic optimization process is as follows:

The first step is to determine the design variables, objective functions, and constraints of the optimization mathematical model; The second step is to build a visual optimization simulation diagram or signal flow diagram based on the system mathematical model; The third step is to build the optimization simulation diagram from the SIMULINK model library by calling the basic module or model component in second steps; The fourth step is to write the optimization M-file, specify the optimization algorithm, and call the MATLAB optimization toolbox subroutine; In the fifth step, the optimization operation is performed, and the results of each iteration are visualized and stored in the optimization database; The sixth step is to find the optimal parameters of the system automatically; The seventh step is the post-processing of the experimental data.

Flowchart of parameter optimization in architectural engineering drawing has been shown in Figure 1.

According to the process mentioned in Section 3.3, the optimization of architectural engineering drawing parameters is realized, and the performance of the building is improved.

4. Experimental Results

In order to verify the optimization effect of this method on the parameters of architectural engineering drawing, the optimization algorithm of architectural engineering drawing parameters based on spark parallel SVM (reference [5] method), the intelligent optimization method of building coverage antenna parameters based on firefly algorithm (reference [6] method), the algorithm based on adaptive determination of DBSCAN algorithm parameters (reference [7] method), and the algorithm based on improved chaos algorithm are adopted The optimization method of architectural engineering drawing parameters (this method) carries out comparative experiments on the optimization precision of architectural engineering drawing parameters, architectural lighting, and optimization time of architectural engineering drawing parameters, and carries out experimental analysis.

After identifying the optimization effect of the proposed method on the parameters of architectural engineering drawing, the results are compared with the adopted algorithm, that is, chaos algorithm in terms of accuracy,
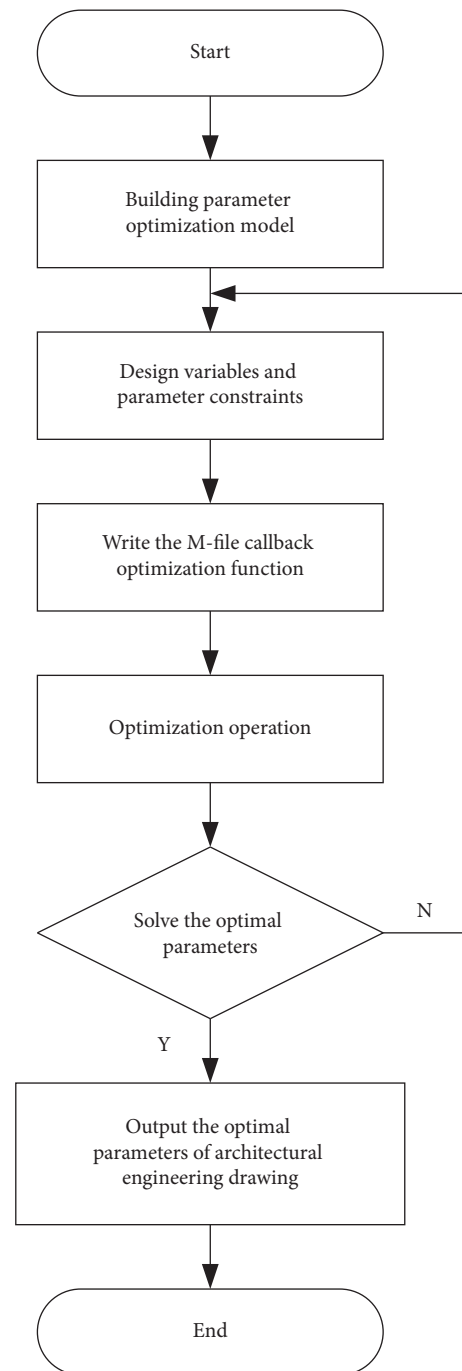
FIGURE 1: Flow chart of parameter optimization in architectural engineering drawing.

architectural daylighting intensity, and time. It has been found that our proposed model outperforms as compared with the methods proposed in research [5–7].

The Comparison of different parameters has been explained in detail as follows:

4.1. Comparison of Parameter Optimization Accuracy in Architectural Engineering Drawing. In order to verify the optimization efficiency of architectural engineering drawing parameters, the methods of literature [5], literature [6],

TABLE 1: Comparison of optimization accuracy of architectural engineering drawing parameters.

Area of structure (m^2)	Optimization precision of architectural engineering drawing parameters (%)			
	Method in literature [5]	Method in literature [6]	Method in literature [7]	Proposed method
50	67	45	65	99
100	68	58	62	99.8
150	65	65	75	95.8
200	58	68	65	99.6
250	53	65	73	97.4
300	63	56	65	97.8
350	68	66	68	98.9
400	72	66	70	97.3
450	78	76	57	98.4
500	74	65	59	97.8
550	73	68	64	95.4
600	74	65	65	99.2
650	79	71	67	99.5
700	76	79	65	98.4
Mean value	69.1	65.2	65.7	98.2

TABLE 2: Architectural daylighting intensity under different methods.

Distance in the building (m)	Daylighting intensity of buildings (lux)			
	Methods in literature [5]	Methods in literature [6]	Methods in literature [7]	Proposed method
1	370	287	353	500
2	230	221	330	487
3	120	112	310	476
4	90	92	276	468
5	50	68	254	456
6	32	45	214	439
7	27	28	187	410
8	12	23	165	387
9	9	12	128	362
10	7	9.4	87	321
11	5	6	33	286
12	3	3.5	26	265
13	2	1.5	21	253
14	1	0.7	8	220
15	0.5	0.6	5.2	165

literature [7], and this paper (proposed method) are used to verify the optimization accuracy of architectural engineering drawing parameters, and the results are shown in Table 1. Comparison of optimization accuracy of architectural engineering drawing parameters has been shown in Table 1.

According to the analysis of Table 1, the optimization accuracy of architectural engineering drawing parameters is different under different methods. When the building area is $100 \, \text{m}^2$, the optimization precision of the method in reference [5] is 68%, the method in reference [6] is 58%, the method in reference [7] is 62%, and the method in this paper is 99.8%. When the building area is $600 \, \text{m}^2$, the optimization accuracy of the method in reference [5] is 74%, that in reference [6] is 65%, that in reference [7] is 65%, and that in this paper is 99.2%. The average values of the optimization accuracy of the parameters in the architectural engineering drawing of the methods of literature [5], literature [6], literature [7], and this paper are 69.1%, 65.2%, 65.7%, and

98.2%, respectively. This method always has the highest optimization accuracy of the parameters in the architectural engineering drawing, which shows that this method can realize the effective optimization of the parameters in the architectural engineering drawing.

4.2. Comparison of Architectural Daylighting. In order to verify the optimization effect of architectural engineering drawing parameters, the actual intensity of architectural lighting under the method of literature [5], literature [6], literature [7], and the method of this paper is compared, and the specific results are shown in Table 2. The architectural daylighting intensity under different methods has been shown in Table 2.

According to Table 2, the daylighting intensity of buildings is different under different methods. When the distance between the two sides of the building is 1 m, the
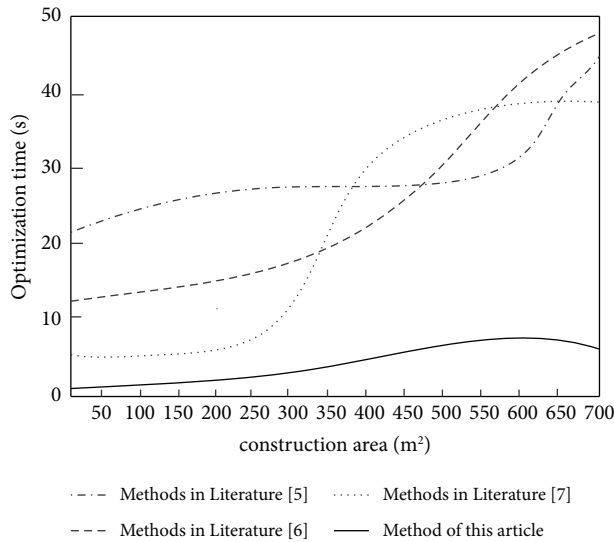
Figure 2: Comparison of optimization time of construction engineering drawing parameters.

daylighting intensity of reference [5] is 370 lux, the daylighting intensity of reference [6] is 287 lux, the daylighting intensity of reference [7] is 353 lux, and the daylighting intensity of the proposed method is 500 lux. When the distance between buildings increases to 10 m, the daylighting intensity of reference [5] is 7 lux, that of reference [6] is 9.4 lux, that of reference [7] is 87 lux, and that of proposed method is 321 lux. This method has higher building daylighting intensity, which shows that the building has higher building daylighting intensity and better daylighting effect.

4.3. Comparison of Optimization Time of Construction Engineering Drawing Parameters. In order to verify the optimization efficiency of architectural engineering drawing parameters, the methods of literature [5], literature [6], literature [7], and this paper are used to detect the optimization time of architectural engineering drawing parameters, and the results are shown in Figure 2.

Analysis of Figure 2 shows that the optimization time of construction engineering drawing parameters is different under different building areas. When the building area is 50 m^2, the optimization time of the method in reference [5] is 24 s, that in reference [6] is 13 s, that in reference [7] is 5 s, and that in this paper is only 1s. When the building area is 400 m^2, the optimization time of building engineering drawing parameters of reference [5] method is 27 s, the optimization time of building engineering drawing parameters of reference [6] method is 21 s, the optimization time of building engineering drawing parameters of reference [7] method is 30 s, and the optimization time of building engineering drawing parameters of the proposed method is only 5 S. When the building area is 600 m^2, the optimization time of the method in reference [5] is 31 s, that in reference [6] is 43 S, that in reference [7] is 40 s, and that in this paper is only 8 s. This method has low optimization time and high optimization efficiency.

5. Conclusion

In this paper, an optimization method of construction engineering drawing parameters based on an improved chaos algorithm is proposed. The optimization model of construction engineering drawing parameters is constructed. The optimization process of construction engineering drawing parameters is designed. The optimization of construction engineering drawing parameters is realized by using an improved chaos algorithm. The modeling strategy of visual parameter optimization environment is constructed. After applying the proposed work to the existing data, it is concluded that proposed method always has the highest optimization accuracy of architectural engineering drawing parameters. The average optimization accuracy of architectural engineering drawing parameters of this method can reach 98.2%. This indicates the efficiency of proposed method in realizing the effective optimization of architectural engineering drawing parameters. Secondly, higher daylighting intensity and better daylighting effect in the buildings are achieved. The daylighting intensity of this method can reach 321 lux when the distance is 10 m. The proposed method shows low optimization time and high optimization efficiency. When the building area is 600 m2, the optimization time is only 8 s. It can be concluded that the proposed improved chaos algorithm outperforms the existing work in terms of accuracy, achieving higher daylighting intensity effect in the buildings as well as reduction in time.

Data Availability

The data used to support the findings of this study are available from the author upon request (maxiaoqiu@jlju.edu.cn).

Conflicts of Interest

The authors declare no conflicts of interest in this article.

Acknowledgments

References

[1] J. Wu, X. Y. Chen, and H. Zhang, "Hyperparameter optimization for machine learning models based on Bayesian optimization," *Journal of Electronic Science and Technology*, vol. 17, no. 01, pp. 26–40, 2019.

[2] N. Nasruddin, S. Sholahudin, P. Satrio, T. M. I. Mahlia, N. Giannetti, and K. Saito, "Optimization of HVAC system energy consumption in a building using artificial neural network and multi-objective genetic algorithm," *Sustainable Energy Technologies and Assessments*, vol. 35, no. 08, pp. 48–57, 2019.

[3] E. Naderi, B. Sajadi, and M. A. Behabadi, "Multi-objective simulation-based optimization of controlled blind specifications to reduce energy consumption, and thermal and visual

discomfort: case studies in Iran," *Building and Environment*, vol. 169, no. 07, pp. 106–110, 2019.

[4] L. Jiang, S. R. Sakhare, and M. Kaur, "Impact of industrial 4.0 on environment along with correlation between economic growth and carbon emissions," *Int J Syst Assur Eng Manag*, vol. 13, no. S1, pp. 415–423, 2021.

[5] J. W. He, L. Z. Liu, and B. Peng, "Research on parameter optimization algorithm of parallel SVM based on spark," *Journal of Wuhan Institute of Technology*, vol. 41, no. 03, pp. 282–289, 2019.

[6] D. L. Dong and P. J. Du, "Intelligent optimization of building coverage antenna based on firefly algorithm," *Designing Techniques of Posts and Telecommunications*, vol. 525, no. 11, pp. 7–11, 2019.

[7] W. J. Li, S. Q. Yan, Y. Jiang, S. Z. Zhang, and C. L. Wang, "Research on method of self-adaptive determination of DBSCAN algorithm parameters," *Computer Engineering and Applications*, vol. 55, no. 05, pp. 1–7, 2019.

[8] N. A. Vasilenko, "Parametric optimization of an architectural object's form as a method to improve its energy efficiency," *IOP Conference Series: Materials Science and Engineering*, vol. 552, no. 1, Article ID 012027, 2019.

[9] K. Anil Kumar and P. S. Chani, "Approaches to parameterization in architectural design," in *Research into Design for Communities, Volume 2. ICoRD 2017: Smart Innovation, Systems and Technologies*, A. Chakrabarti and D. Chakrabarti, Eds., Springer, Singapore, 2017.

[10] G. Canestrino, "Considerations on optimization as an architectural design tool," *Nexus Network Journal*, vol. 23, no. 4, pp. 919–931, 2021.

[11] E. Naderi, B. Sajadi, M. A. Behabadi, and E. Naderi, "Multi-objective simulation-based optimization of controlled blind specifications to reduce energy consumption, and thermal and visual discomfort: case studies in Iran," *Building and Environment*, vol. 169, no. 02, Article ID 106570, 2020.

[12] R. Paul and S. K. Dalui, "Optimization of alongwind and crosswind force coefficients on a tall building with horizontal limbs using surrogate modeling," *The Structural Design of Tall and Special Buildings*, vol. 30, no. 4, pp. 24–32, 2020.

[13] X. Chen, J. Huang, and H. Yang, "Multi-criterion optimization of integrated photovoltaic facade with inter-building effects in diverse neighborhood densities," *Journal of Cleaner Production*, vol. 248, no. 05, Article ID 119269, 2020.

[14] A. Jadhav, M. Kaur, and F. Akter, "Evolution of software development effort and cost estimation techniques: five decades study using automated text mining approach." *Mathematical Problems in Engineering*, vol. 2022, pp. 1–17, Article ID 5782587, 2022.

[15] R. N. Cai, Z. G. Sun, H. Yu, E. L. Meng, J. Wang, and M. Dai, "Review on optimization of phase change parameters in phase change material building envelopes," *Journal of Building Engineering*, vol. 35, no. 16, Article ID 101979, 2021.

[16] M. Kaur, A. Jadhav, and F. Akter, "Resource selection from edge-cloud for IIoT and blockchain-based applications in industry 4.0/5.0,"," *Security and Communication Networks*, vol. 2022, Article ID 9314052, 10 pages, 2022.

Hindawi

Retraction

Retracted: A Plain Bayesian Algorithm-Based Method for Predicting the Mental Health Status and Biomedical Diagnosis of University Students

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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Wang, "A Plain Bayesian Algorithm-Based Method for Predicting the Mental Health Status and Biomedical Diagnosis of University Students," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2617488, 5 pages, 2022.

Research Article

A Plain Bayesian Algorithm-Based Method for Predicting the Mental Health Status and Biomedical Diagnosis of University Students

Jiao Wang [ID]

Center for Ideological and Political Education & Guidance Center for Student Psychological Development, Northeast Normal University, Changchun 130024, China

Correspondence should be addressed to Jiao Wang; 19401151@masu.edu.cn

The purpose of this study was to assess e-learning during Corona epidemic regarding advantages, limitations, and their recommendations for managing learning during the epidemic. Based on a case study, this study used qualitative research. Sixteen students from King Saud University's College of Education were invited to take part. These students receive their online lectures via the "Zoom" application. A 20-minute WhatsApp one-on-one semiorganized interview was likewise utilized. To guarantee the reliability, iCloud was utilized to record gatherings and meetings for direct record (adaptability, constancy, confirmability, and validity). Results were presented in three themes: advantages of employing distance education, limitations of usages, and recommendations for improvements. Analyzing the feedbacks collected from students by the four interviewers, important characteristics of distance education emerged. They were student-centered learning, which included: comfortable, self-directed learning, asynchronous learning, and flexibility. The most common limitations associated with distance education, in general, included inefficiency, that is, lack of student feedback, and lack of attentiveness. As for recommendations for improvements the most obvious characteristics that became evident in students' responses were teaching and assessment and quality enhancement.

1. Introduction

Since the beginning of 2020, the world has witnessed exceptional circumstances associated with the prevalence of the new Corona virus or Covid-19 [1]. Educational institutions have been affected by these decisions, and it is no longer possible to continue the educational process in its traditional form, and millions of students and their teachers at all educational stages have been kept in their homes. The form of the spread of this virus or the Corona epidemic is a major crisis that all countries of the world faced without exception, and the countries of the world rushed to search for alternative methods to prevent the educational process from stopping, and not all countries were at the same level in dealing with this emergency situation. Different experiences emerged to deal with this emergency, and this was followed by the emergence. The term emergency remote education is along with the term

e-learning, and UN institutions such as UNESCO [2] and other institutions have intervened to follow this emergency matter and provide a helping hand to various countries of the world, given that this epidemic has formed a common denominator for the suffering of millions of people worldwide, and part of that suffering is due. To the fears associated with the spread of the epidemic and the number of daily deaths, and the denial of the practice of normal humanitarian activity.

1.1. Saudi Arabia Experiences in Applying E-Learning during Corona Epidemic. The Kingdom of Saudi Arabia has a rich experience in managing crises facing the educational process. This appeared at the southern border. In 2016, a distance education program was signed, in cooperation between the Ministry of Education and the Education Development Company, and the Minister of Education stated

at the time that providing this quality service in the educational field is a duty to enable students to continue their learning under any circumstances and that the Education Development Company has prepared for this program and has the capabilities to fully implement it to help provide education and service at this time, especially for those affected by crises, indicating that the Ministry of Education provided educational alternatives in the southern border and benefited from developing electronic programs and broadcasting some programs through TV channels prepared for this purpose, and he confirmed the Ministry's readiness to provide assistance at any time in all circumstances [3, 4].

As for the management of the crisis that resulted from the Corona epidemic, the Saudi Ministry of Education announced the closure of various educational institutions since the eighth of March in accordance with the Noble Decree No. (42874), and a specialized committee was formed in the Ministry of Education to follow up the developments of the spread of the Corona epidemic, and the Ministry took the initiative to define a variety of packages.

As per the Saudi Ministry of Education, e-learning stages for public and private schooling have been actuated, notwithstanding the accompanying critical measures.

The virtual school building was finished in multi week, during which the school's central command and satellite telecom were outfitted, 20 savvy sheets were provided and introduced, and the school personnel was prepared to photo educational clasps.

With the assistance of 276 educators and 73 bosses, everyday examples were ready to make sense of the educational program. 3368 examples were made sense of, adding up to 1684 hours of educating time.

In addition to the curriculum review lessons, 123 teachers and 73 supervisors participated. 1107 lessons were reviewed, totaling 554 hours of teaching time.

The Ministry carried out distance education involving the techniques for synchronous collaboration and nonconcurrent communication as follows.

Concurrent connection was accomplished through the virtual school application, which incorporated the bound together education framework, the eye entryway, the door representing things to come, and the virtual kindergarten application. Nonconcurrent connection was utilized by means of Ain satellite stations and Lessons Ain stations on YouTube. The past stations and stages gave numerous choices to male and female students to proceed with education and advancing from a distance, by means of the Internet or through satellite stations, for the people who do not approach the Internet, and these stations saw a ton of collaborations among male and female students and contributed extraordinarily to the fruitful continuation of the educational cycle. Following is a short outline of these channels.

Future Gate is a coordinated e-learning stage for students, all things considered. It has received over 20 million visits, and its services have benefited over 700,000 students.

Ain Satellite Channels are an assortment of 20 satellite stations that cover every single educational level. The channels are functional 24 hours a day, seven days a week.

Ain Channels on YouTube: These are live YouTube channels with more than 60 million perspectives.

Ain (National Education Portal): The gateway offers advanced content that incorporates more than 45,000 virtual educational materials and north of 2000 computerized books.

Virtual Kindergarten is a thorough educational framework for the youth stage.

The framework contains advanced educational illustrations and exercises in an intelligent educational climate between the educator and the understudy. The Ain stage gives day-to-day clarifications of the educational plans for every educational stage, and the Kingdom's choice to enact the "Board" framework came following the spread of the Corona infection, and the framework gives many advantages to instructors, including the following:

Keeping track of student attendance and absences.

Students can be given assignments, tests, and courses.

Create virtual lessons.

Make educational ways for a select gathering of students.

Answer understudy requests sent through educator rooms.

The brought-together education framework site has become one of the most noticeable arrangements that anyone could hope to find for distance education in the Kingdom of Saudi Arabia to speak with students, and it remembers all administrations connected with education for Saudi Arabia for all male and female students. This short show of the Kingdom of Saudi Arabia's experience shows that there are numerous choices accessible to male and female students in the Kingdom to follow the educational cycle and to furnish distance education with various options through the Internet or satellite stations, as well as the Ministry of Education's nearby drive to go to the vital lengths and not stop educational administrations for all degrees of education. The Saudi experience concurred with its partner in the United Arab Emirates, as the UAE has broad experience and mastery in the field of involving innovation in education, and this experience varied from past encounters, which saw a stoppage in the use of distance education methods or a more prominent accentuation on the utilization of the Internet around here.

The Saudi Ministry of Education's drive to manage the Corona plague, and the danger it postures to the educational cycle can be made sense of considering the Kingdom's past involvement with overseeing educational emergencies, as the Saudi Ministry of Education had past rich involvement with giving distance education administrations to male and female students on the Kingdom's southern lines, and the arrangement included distance education administrations. It safeguarded the understudy's principal privileges to education by placing him in direct contact with his educators through the school twinning venture. Other electronic educational options, for example, the "illustrations" educational stations and the "AAli" satellite station, were likewise enacted, with the stages of online education benefiting.

2. Objective

The goal of this study was to evaluate e-learning during Corona epidemic regarding advantages, limitations, and their recommendations for managing learning during the epidemic.

3. Method

3.1. Design. This study utilized qualitative case study research, in which the scientist examines a reality, contemporary limited framework or various limited frameworks over the long run by social event definite, top to bottom data and detailing case subjects and case depictions [5].

3.2. Participants and Procedures. Sixteen students from King Saud University's College of Education were invited to take part. These students received their online lectures via the "Zoom" application. Furthermore, a 20-minute WhatsApp one-on-one semiorganized interview was utilized [6]. To guarantee dependability, iCloud was utilized to record gatherings and meetings for direct record (transferability, dependability, confirmability, and credibility). Four interviewers; the author besides three other lecturers, conducted the interviews. The data were thematically analyzed: compiling, disassembling, reassembling, and interpretation by all the interviewers each on his own, and then the data were corroborated to ensure analytical accuracy [5].

4. Results

4.1. Presentation of Findings. Results were presented in three themes: Advantages of employing distance education, limitations of usages, and recommendations for improvements.

Theme 1. Advantages of employing distance education. Analyzing the feedbacks collected from students by the four interviewers, an important characteristic of distance education emerged. It is student-centered learning, which includes comfortable, self-directed learning, asynchronous learning, and flexibility. For example, the majority of the students agreed with Student 18, who stated, "*It is easy and more comfortable for me now to listen to the lecture while sitting in my bedroom, or anywhere that I get used to and feel comfortable*" [7]. This new learning climate focused on electronic organizations has empowered college students to get individualized help along with learning plans that are more reasonable to them and separate from other students. Obviously, e-learning can empower students at higher educational levels to get their education while likewise seeking after their own objectives and keeping up with their own professions, without sticking to an inflexible timetable [8].

As for Self-directed Learning, a large portion of the understudies concurred with Student 3 who placed, "*I think Distance Education is making good students more active and self-learner, that is, it helps them in becoming engaged with growing experiences like securing data, arranging, and assessing learning exercises.*" All in all, independent students ordinarily take part more effectively in learning errands, for example, perusing web-based learning material, finishing homeroom assignments, and arranging and assessing learning achievements [9]. An independent student effectively takes part in growing experiences like data securing, arranging, and assessing learning exercises. Dynamic learning systems can help understudy investment while additionally further developing learning and execution [10, 11].

Concerning asynchronous learning, a large portion of the understudies concurred with Student 7 who set, "*One can download lectures at any time. It is easy for us to go back and go through the whole video for revision.*" Asynchronous learning, in other words, assists in making learning self-paced, independent, and, most importantly, student-centered [12]. Accordingly, this sort of learning is probably going to framework students' earlier information with new ideas [13], it allows for peer group discussions, and it promotes decisive reasoning and profound learning [14].

As for flexibility, the majority of the students agreed with Student 18, who proposed, "*Although we are away from our university, and cannot attend lectures, however, distant learning during COVID 19 enables us to continue our education and have all our lectures without discontinuity.*" As such, students have more opportunity in choosing when, how, and with what content and exercises they take part in web-based learning situations [15]. This versatility requires students checking and changing their way of behaving and activities comparable to the particular learning setting (Zimmerman, 2000).

Theme 2. Limitations of usages. Analyzing the feedbacks collected from students by the four interviewers, students reported some Limitations of usages, concerning distance learning. The most obvious characteristic that became evident in student's responses was inefficiency, which included lack of student feedback and lack of attentiveness. Concerning lack of student feedback, the vast majority of the understudies concurred with Student 2 who set, "*the professor does not know for sure whether his students are getting the point or not. That is, he does not care much of getting students feedback. He only gives the lecture as if it were a chore to do.*" In other words, due to the epidemic, higher institutions in every corner of the world, including my country as well, are compelled to suspend up close and personal classes. That is, they have rushed to remote teaching and online classes, posing unprecedented challenges in terms of learning new technologies and gaining access to necessary facilities such as laboratories [16].

Concerning lack of attentiveness the majority of the understudies concurred with Student 15 who set, "*I think that type to instruction has some limitations such as lacking students attention . why? Students know that they will get the recordings at any time, so they might not listen the lecture properly.*" In other words, in a real-world classroom, a teacher easily captures students' attention and adapts lessons accordingly to maximise student interest and participation [17]. However, in distance education, students feel free to attend or not attend to the lecturer, as they think they will get the recordings at any time.

Theme 3. Recommendations for improvements. Analyzing the feedbacks collected from students by the four interviewers, students reported some recommendations for improvements, concerning distance learning. The most obvious characteristics that became evident in students responses were teaching and assessment and quality enhancement. Teaching and assessment included online examination. As for online examination, the majority of the understudies concurred with Student 6 who set, "*online examination helps university teachers to reach remote students.*" Another student (Student 3) proposed that "*the online examination may result in an equal level of student performance.*" Another student (Student 17) stated that "*online examination helps to boost students' results.*" As such, students favor electronic appraisals since they have more control, easy to use points of interaction, and tests that look like learning conditions and sporting exercises. When contrasted with paper tests, e-appraisal gives quick criticism, which supports the improvement of learning levels. This is consistent with the findings of [18–20].

As for quality enhancement, majority of the students agreed with Student 1 who stated, "*Being out of sight of their instructors, will allow students to collaborate together, thus writing the same assignments from the same source. I think, there should be plagiarism software to check assignment.*" Some researchers blame the Internet for the increased "opportunities" for cheating [21], while others believe the increased incidence of cheating is due, at least in part, to the use of Internet [22].

5. Discussion

The goal of this study was to evaluate e-learning during the Corona epidemic regarding advantages, limitations, and their recommendations for managing learning during the epidemic. Results were presented in three themes: advantages of employing distance education, aimitations of usages, and recommendations for improvements. Analyzing the feedbacks collected from students by the four interviewers, important characteristics of distance education emerged. They were Student-centered learning, which included: comfortable, self-directed learning, asynchronous learning, and flexibility. This is consistent with findings, which determined that flexible learning, defined as a learner-centered approach to learning, provided students with different learning options, which in turn helps to make learning outcomes useful and exciting [23–26]. During the COVID-19 pandemic, we require enough time to extend and sustain online education.

The most common limitations associated with distance education in general included the inefficiency, that is, lack of student feedback and lack of attentiveness. In the traditional classroom, students meet each other in the space; exchange ideas, information, and knowledge; ask the professor; interact with them; and receive feedback from colleagues and teachers, while sharing ideas, knowledge, and information in real time in the virtual environment, i.e., in distance learning, is considered missing to a large extent. Concerning reported limitations, this result, in the same line, indicates it

is absurd to teach students through distance learning without preparing the climate, environment, and students for this type of learning [27]. Teachers should be skilled, on one hand, and students should be prepared for this, on the other hand.

With the advent of social media, students now have instant access to a wealth of information. Internet search results are also electronic, and students can easily copy and paste the information into assignments, papers, and other documents [28].

6. Conclusion

This paper points to assess e-learning during the Corona epidemic regarding advantages, limitations, and their recommendations for managing learning during the epidemic. Results were presented in three themes: advantages of employing distance education, limitations of usages, and recommendations for improvements. Analyzing the feedbacks collected from students by the four interviewers, important characteristics of distance education emerged. They were Student-centered learning, which included comfortable, self-directed learning, asynchronous learning, and flexibility, inefficiency, which included ack of student feedback, and lack of attentiveness, and teaching and assessment and quality enhancement. Teaching and assessment included online examination and Quality enhancement.

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] A. Lee, "Wuhan novel coronavirus(COVID-19):why global control is challenging?" *Public Health*, vol. 179, pp. A1–A2, 2020.

[2] Unesco, "Education For Sustainable Development Goals: Learning Objectives," 2020, https://unesdoc.unesco.org/ark:/48223/pf0000247444.

[3] M. Mahrous, "Establishing A contemporary educational theory for the management of the novel Corona virus (COVID - 19)," *The educational Journal*, vol. 5, no. 3, pp. 465–500, 2020.

[4] M. B. Alazzam, H. Mansour, F. Alassery, and A. Almulihi, "Machine learning implementation of a diabetic patient monitoring system using interactive E-app," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–7, 2021.

[5] J. Creswell and C. N. Poth, *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, SAGE Publications, inc, California, CA, USA, 2017.

[6] J. Creswell, *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, SAGE Publications, inc, California, CA, USA, 3rd edition, 2014.

Research Article

Supervised Approach to Identify Autism Spectrum Neurological Disorder via Label Distribution Learning

N. V. L. M Krishna Munagala [ID],[1] V. Saravanan [ID],[2] Firas Husham Almukhtar [ID],[3] Naveed Jhamat [ID],[4] Nadeem Kafi [ID],[5] and Samiullah Khan [ID][6]

[1]*Department of Electrical Electronics and Communication Engineering, GITAM Institute of Technology,*
 GITAM Deemed University, Visakhapatnam, Andhra Pradesh 530045, India
[2]*Dambi Dollo University, Dambi Dollo, Ethiopia*
[3]*Department of Computer Technical Engineering, Imam Ja'afar Al-Sadiq University, Kirkuk, Iraq*
[4]*Department of Information Technology, University of the Punjab, Gujranwala Campus, Gujranwala, Pakistan*
[5]*Department of Computer Science, National University of Computer and Emerging Sciences, Karachi, Pakistan*
[6]*Department of Maths, Stats & Computer Science, The University of Agriculture Peshawar, Peshawar, KP, Pakistan*

Correspondence should be addressed to V. Saravanan; saravanan@dadu.edu.et

Autism Spectrum Disorder (ASD) is a complicated collection of neurodevelopmental illnesses characterized by a variety of developmental defects. It is a binary classification system that cannot cope with reality. Furthermore, ASD, data label noise, high dimension, and data distribution imbalance have all hampered the existing classification algorithms. As a result, a new ASD was proposed. This strategy employs label distribution learning (LDL) to deal with label noise and uses support vector regression (SVR) to deal with sample imbalance. The experimental results show that the proposed method balances the effects of majority and minority classes on outcomes. It can effectively deal with imbalanced data in ASD diagnosis, and it can help with ASD diagnosis. This study presents a cost-sensitive approach to correct sample imbalance and uses a support vector regression (SVR)-based method to remove label noise. The label distribution learning approach overcomes high-dimensional feature classification issues by mapping samples to the feature space and then diagnosing multiclass ASD. This technique outperforms previous methods in terms of classification performance and accuracy, as well as resolving the issue of unbalanced data in ASD diagnosis.

1. Introduction

Autism spectrum disorder (ASD) is a series of complex neurodevelopmental disorders, and its clinical manifestations are mainly social interaction disorders, verbal communication disorders, and stereotyped repetitive movements [1, 2]. Statistics from the US Centers for Disease Control and Prevention show that the prevalence of autism in American children is as high as 1 : 59. This shows that autism has become a rather serious health problem and there is an urgent need to develop an effective method for timely diagnosis. However, because the physiological cause of autism is not clear, medical diagnosis can only be based on the patient's symptoms and

feedback, qualitative/quantitative testing information, and the physician's personal experience, which has great uncertainty [3]. Therefore, it is of great significance to use computers to assist in the diagnosis of autism.

Studies have shown that autism spectrum disorders are related to the abnormal brain function in patients and resting-state functional magnetic resonance imaging, which reflects functional changes such as brain metabolic activity in patients under a resting state, is reflected using blood oxygen-dependent levels [4, 5]. Resonance image (resting-state functional magnetic resonance imaging, rs-fMRI) has become a powerful tool for quantifying neural activity in the brain and has gradually become one of the important means for the

study of brain diseases such as ASD [6, 7]. Based on this diagnosis, researchers have proposed a variety of computer-aided autism diagnosis algorithms [8, 9]. For example, the authors used high-order functional connectivity matrix for auxiliary diagnosis of autism. He proposed multivariate graph learning for auxiliary diagnosis of autism, the authors explored the relationship between brain regions through deep learning and Correlation for auxiliary diagnosis of autism and so on [10]. However, these methods can only deal with dichotomous problems, and in clinical practice, autism spectrum disorder includes several disorders related to developmental disorders, such as autism [11] and Asperger's syndrome (Asperger's disorder), nonspecific general developmental disorders (pervasive developmental disorder not otherwise specified, PDD-NOS), and so on. Most of the existing auxiliary diagnosis models for autism can only solve the problem of binary classification and cannot distinguish several related diseases of ASD at the same time. In addition, these methods also do not deal with label noise in a targeted manner [12]. Labeling noise is a challenge involved in the auxiliary diagnosis of multiclass ASD and has serious adverse effects on classifier performance [13]. Label noise refers to the deviation between the target label of the training sample and the true label of the corresponding instance. There are many factors in the generation of labeling noise, such as the subjectivity of the labeling process, the low recognizability of the samples to be labeled, and communication/coding problems. Labeling noise is prevalent in autism diagnosis scenarios. Subjectivity in the diagnostic process, inconsistent diagnostic criteria, and blurring of the boundaries of ASD subcategories contribute to labeling noise [14].

The class imbalance problem under high-dimensional features is another challenge involved in the auxiliary diagnosis of multiclass ASD [15]. The neuroimaging data usually used for the auxiliary diagnosis of ASD often have hundreds or thousands of features, and the number of training samples is very limited, which may easily lead to overfitting problems during classifier training. Moreover, the samples used to construct the ASD classifier have the problem of class imbalance, which causes the classification prediction results to be biased towards the majority class [16, 17]. This paper proposes a cost-sensitive label distribution support vector regression learning for auxiliary diagnosis of ASD [18]. First of all, multiclass ASD auxiliary diagnosis is faced with the problem of label noise, and the unique label form of label distribution can better overcome the influence of label noise on the classifier through the description of the same sample by different labels to accurately express the difference between labels. The degree of correlation makes the learning process contain richer semantic information, can better distinguish the relative importance of multiple markers, and has better pertinence to the problem of marker noise in the auxiliary diagnosis of ASD [19, 20]. The kernel approach is introduced at the same time as the support vector regression. The linearly inseparable data in the original input space may be transferred into a linearly separable feature space using the kernel method's nonlinear mapping, offering additional discriminative information. Finally, a cost-sensitive technique is devised to address the issue of category imbalance. The

algorithm may adjust to the demands of actual applications to some degree and treat a limited number of individuals equitably by introducing the imbalance of misjudgment costs of various categories in reality.

Label distribution learning (LDL) is designed to cope with label noise in this technique, while support vector regression (SVR) is also used to handle the sample imbalance. According to the results of the trials, the proposed technique optimizes the effects of majority and minority classes on outcomes. It can handle skewed data in ASD diagnosis and can assist with ASD diagnosis. This study provided a cost-sensitive technique for correcting sample imbalance using a support vector regression (SVR)-based method to reduce label noise. The label distribution learning approach addresses high-dimensional feature classification challenges by mapping data to the feature space and then diagnosing multiclass ASD. In terms of classification performance and accuracy, our proposed strategy outperforms earlier methods, as well as eliminates the challenge of unbalanced data in ASD diagnosis.

However, the improved model is still biased towards the majority class to some extent, and the imbalanced data problem should be improved further as a future study. Researchers can further try to improve the data sampling method or use the synthetic minority sample method as future prespective

1.1. Organization. The paper is framed into several sections where Section 1 states about the Introduction followed by related work section in Section 2. Section 3 states about cost-sensitive marker distribution learning for ASD-aided diagnosis, followed by Section 4 that describes the evaluation of proposed methodology. The final section is the concluding section numbered 5 that discusses the results obtained in the study.

2. Related Work

2.1. Labeled Distribution Learning. Label distribution learning (LDL) is a machine learning method that has emerged in recent years [21]. It introduces the concept of label distribution on the basis of single-label and multilabel learning [22, 23]. In a multimarket scenario, if a sample is related to multiple markers, the importance of these markers to the sample will generally be different, and the marker distribution is a marker form that describes the importance of different markers to the same sample. Label distribution learning is a machine learning method that takes label distribution as the learning target and has been applied in many fields. Author proposed a deep label distribution learning algorithm combining convolutional neural network and label distribution learning to estimate age by face, and Author uses wheel of emotions to automatically identify the user's emotional state from the text. Author proposed an algorithm based on multivariate label distribution to detect head pose [24, 25]. However, it has not yet been reported for the auxiliary diagnosis of brain diseases. This study aimed to identify particular qualities that aid in the automation of the diagnostics, as well as evaluating and contrasting various
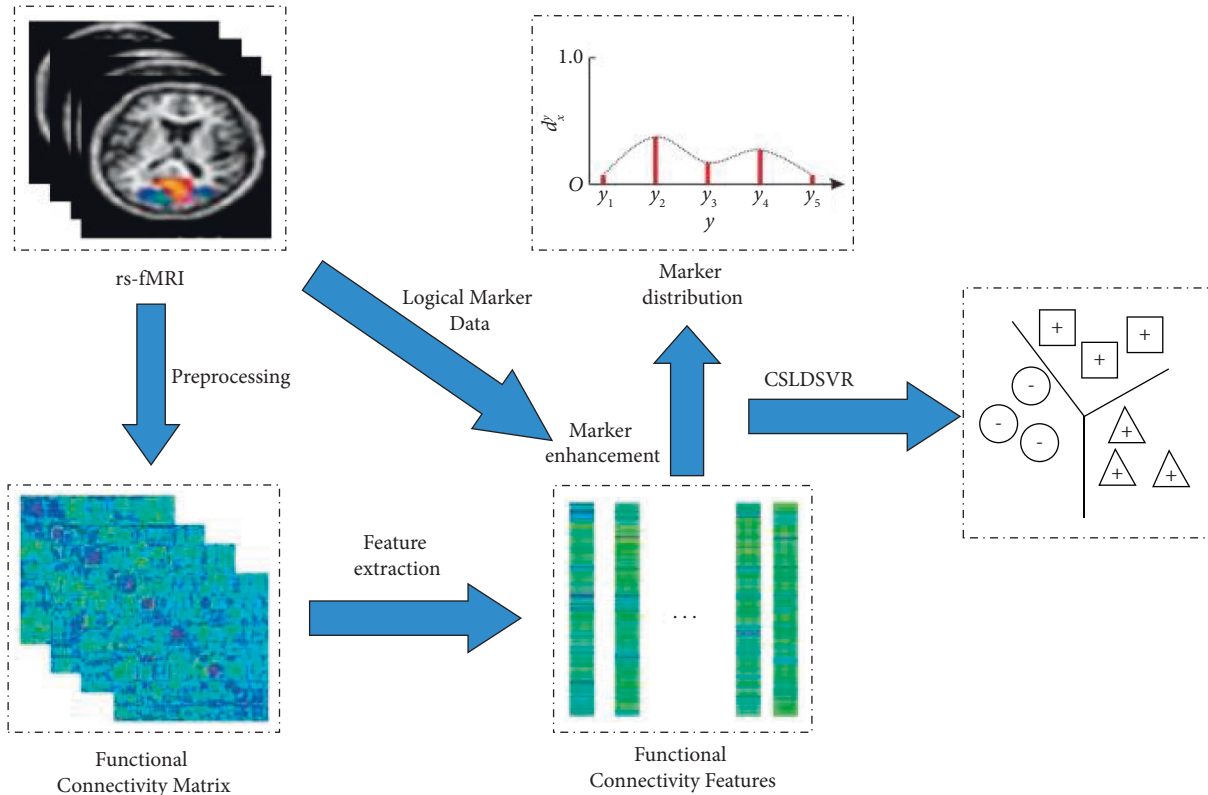
FIGURE 1: Proposed label distribution support vector regression for cost sensitivity.

machine learning techniques [26]. The functional connectivity structure acquired from resting-state MRI that was being used to construct the auto-encoder that is semi-supervised for autism diagnosis in this research is proposed [27].

2.2. Marker Enhancements. Label distribution learning requires that the training data contain label distribution information. However, in real life, people often label samples in the form of single-label or multilabel, making it difficult to directly obtain label distribution information. Nonetheless, the labels of these data still contain relevant information about the distribution of the labels. Marker enhancement enhances the supervised information of samples through the implicit correlation between different sample markers, thereby achieving better results in marker distribution learning [28]. For example, the authors proposed tag augmentation as an auxiliary algorithm for tag distribution learning, which is used to mine the implied tag importance information in the training set, promote the original logical tag to tag distribution, and assist tag distribution learning. The authors proposed label-enhanced multilabel learning to reconstruct latent label importance information from logical labels to improve the performance of label distribution learning [29].

3. Cost-Sensitive Marker Distribution Learning for ASD-Aided Diagnosis

3.1. Symbolic Representation. The main symbols in this paper are expressed as follows: Use $x_i \in R^q$ to represent the ith

sample, where q represents the dimension of the feature vector; $X = [x_1, \ x_2, \cdots, x_N] \in R^{q \times N}$; $l_i = [l_i^1, l_i^2, \ldots, l_i^K,]^T$ represents the logical token corresponding to x_i, where K represents the number of possible tokens; and $l_i^j \in \{0, 1\}$. Similarly, $d_i = [d_i^1, d, \ldots, d_i^K,]^T \in R^K$ represents the label distribution of the ith sample, where $d_i^j \in [0, 1]$ represents the jth value of the label distribution of the ith sample, satisfying $\sum_{j=1}^K d_i^j = 1, D = [d_1, d_2, \cdots, d_N] \in R^{K \times N}$.

3.2. Proposed Methodology. The label distribution learning algorithm for multiclass autism auxiliary diagnosis proposed in this paper is shown in Figure 1. First, the rs-fMRI images are preprocessed, and the functional connectivity matrix is constructed on this basis, and the functional connectivity feature vector of each sample is obtained based on the functional connectivity matrix. At the same time, combining the logical marker data and functional connectivity features for marker enhancement, the marker distribution form of the sample is obtained. Finally, a cost-sensitive label distribution learning model is carried out to obtain a multi classification model for the auxiliary diagnosis of autism.

3.3. Marker Distribution Mechanism. Label distribution learning describes the degree of correlation between each label and sample by introducing a descriptive degree, so it can obtain richer semantic information from the data than multilabel and more accurately express the relative importance difference of multiple labels of the same sample. However, the basic requirement of labeled distribution

TABLE 1: Evaluation measures.

	Index	Formula
Mark distribution metrics	Chebyshev↓	$\mathrm{Dis}_1 = \max\limits_{j \in [1,K]} |d_i^j - \widehat{d}_i^j|$
	KL↓	$\mathrm{Dis}_2 = \sum_{j=1}^K d_i^j \ln d_i^j / \widehat{d}_i^j$
	Clark↓	$\mathrm{Dis}_3 = \sqrt{\sum_{j=1}^K (d_i^j - \widehat{d}_i^j)^2 / (d_i^j + \widehat{d}_i^j)^2}$
	Canberra↓	$\mathrm{Dis}_4 = \sum_{j=1}^K |d_i^j - \widehat{d}_i^j| / |d_i^j| + |\widehat{d}_i^j|$
	Intersection↑	$\mathrm{Sim}_1 = \sum_{j=1}^K \min(d_i^j, \widehat{d}_i^j)$
	Cosine↑	$\mathrm{Sim}_2 = d_i \cdot \widehat{d}_i / |d_i| \cdot |\widehat{d}_i|$
Multiclass metrics	Precision	$P = 1/N \sum_{i=1}^N \mathrm{xnor}(l_i, \widehat{l}_i)$
	mAB	$m\mathrm{AB} = 1/N \sum_{i=1}^N p_j$

learning is to have labeled distributed datasets, which is often difficult to meet in reality. The marker distribution data can be obtained by transforming a given multimarket form sample by a marker enhancement method. The label enhancement method based on FCM (fuzzy C-means) and fuzzy operation is adopted [30]. The basic idea is as follows:

(1) Use FCM to divide N samples into p fuzzy clusters, and find the center of each cluster, so that the sum of the weighted distances from all training samples to the cluster center is the smallest. Equation (1) lists the specific weighted distance formula:

$$m_{x_i}^k = \frac{1}{\sum_{j=1}^P \left(\mathrm{Dist}(x_i, \mu_k)/\mathrm{Dist}(x_i, \mu_j)\right)^{1/\beta-1}}. \quad (1)$$

Among them, $m_{x_i}^k$ represents the membership degree of the ith sample to the kth cluster center, μ_k represents the kth cluster center, β is a fuzzy factor greater than 1, Dist $(*, *)$ represents the distance measure, and each sample the membership degree represents the strength of the association between the sample and the cluster. The clustering result of traditional FCM is greatly affected by the initial value and cannot ensure convergence to the global optimal solution, but in label enhancement, the clustering result of FCM is only used as a transitional bridge. Although the clustering result fluctuates, however, it has little effect on the results of label enhancement, and the gaps between the Chebyshev distance and the KL divergence (Kullback–Leibler divergence) of the results of multiple label enhancements are both below 10^{-6}.

(2) Construct an association matrix A between markers and clusters. The elements in the matrix represent the degree of association between markers and clusters. The calculation method of the association matrix is as follows:

$$A_j = A_j + m_{x_i}^k, \text{ if } l_i^j = 1. \quad (2)$$

In the formula, A_j is the jth row of the matrix and A_j is the sum of the membership degree vectors of the

samples of the jth class. After the rows are normalized, the association matrix A can be regarded as a fuzzy relationship matrix of clustering and labeling.

(3) According to the fuzzy logic reasoning mechanism, the fuzzy synthesis operation is performed on the association matrix and the membership degree, and the membership degree of the sample to the label is obtained [31]. After normalization, it is the label distribution.

The marker enhancement based on FCM and fuzzy operation introduces cluster analysis as a bridge. Through the compound operation between the membership degree of the sample to the cluster and the membership degree of the cluster to the marker, the membership degree of the sample to the marker, that is, the marker, is obtained distributed. In this process, the topological relationship of the sample space is mined through fuzzy clustering, and this relationship is projected to the label space through the association matrix, so that the simple logical labeling generates richer semantic information and transforms it into a label distribution.

4. Evaluation of Proposed Methodology

4.1. Evaluation Metrics. This paper uses both the evaluation metric of the label distribution and the evaluation metric of the multiclassification task for algorithm evaluation. All evaluation indicators and calculation formulas are shown in Table 1. The first six are evaluation indicators for labeled distribution learning, and the last two are evaluation indicators for multiclassification tasks. "↑" after the index name means that the larger the value, the better the algorithm effect; with "↓," the smaller the value, the better the algorithm effect.

In Table 1, P_j is the precision of the jth class, xnor is the XOR calculation, Dis is the distance, Sim is the similarity, and mAP is the macro-averaging precision.

4.2. Dataset Used. All rs-fMRI datasets used in this paper were obtained from the ABIDE website (Autism Brain Imaging Data Exchange, http://fcon_1000.projects.nitrc.org/indi/abide/). Table 2 shows the composition of each type of sample in each dataset. Taking the NYU (New York University) dataset as an

TABLE 2: Statistics of datasets.

Dataset	Number of samples	Normal	Autism	Asperger's syndrome
NYU	175	104	56	24
UM	140	73	54	16
KKI	52	37	7	8
Leuven	104	64	23	17
UCLA	83	51	18	14

TABLE 3: Comparison algorithms.

Comparison algorithm name	Description of the comparison algorithm
PT-SVM	Based on the problem-transformed SVM
PT-BAYES	BAYES-based gauss distribution
AA-KNN	The algorithm-based KNN
AA-BP	The algorithm-based BP neural network uses the softmax activation output as the predicted label distribution
SA-IIS	IIS based on dedicated algorithm uses an improved iterative scaling algorithm to optimize the objective function
LDSVR	LDSVR based on a dedicated algorithm
Decision tree	An instance-based inductive learning method
KNN	An instance-based classification method

example, the data collection institution of the NYU dataset is New York University. During the collection process, the subjects remained in a still state and did not perform any actions. The specific parameters are shown in Table 2.

In Table 2, UM stands for the University of Michigan, KKI for the Kennedy Krieger Institute, Leuven for the University of Leuven, and UCLA for the University of California, Los Angeles.

Although brain regions are spatially isolated from each other, the neural activity between them influences each other. This paper uses the brain functional connectivity matrix between brain regions as a classification feature [32]. The calculation step (preprocessing step) of the functional connectivity matrix is as follows:

(1) According to the resting-state functional magnetic resonance imaging data, use the DPARSF (data processing assistant for resting-state fMRI) tool to extract the average time-series signals of each brain region, calculate the Pearson coefficient between the brain regions, and obtain the functional connectivity matrix

(2) Take each row of the functional connectivity matrix as the feature description of each brain region, take the upper triangular matrix of the functional connectivity matrix, and connect the rows in series to obtain the corresponding eigenvectors

4.3. Proposed Algorithm.
The proposed CSLDSVR method is compared with six existing LDL algorithms and two multiclassification algorithms. Two multiclassification algorithms are decision tree and K-nearest neighbor (KNN), both of which are classic multiclassification algorithms [33, 34]. The six existing LDL algorithms are PT-SVM, PT-BAYES, AA-KNN, AA-BP (back propagation), SA-IIS (improved iterative scaling), and LDSVR, where "PT" stands for problem

TABLE 4: Range of parameters.

Parameter name	Parameter range
Weight factor	0.001, 0.01, 0.1, 1, 10, 100, 1000
Type of kernel function	Linear kernel, polynomial kernel, Gaussian kernel
Insensitive area size	0.0001, 0.001, 0.01, 0.1
The kernel bandwidth of the Gaussian kernel	0.01, 0.1, 1, 10, 100

transformation, "AA" for algorithm adaptation, and "SA" for specialized algorithm [35, 36]. The specific description of the comparison algorithm is shown in Table 3.

The CSLDSVR algorithm proposed in this paper has four parameters, namely, the weight coefficient C, the type of kernel function, the size of the insensitive region ε, and the kernel bandwidth of the Gaussian kernel. The specific range of parameters is shown in Table 4. The results were calculated using ten-fold cross-validation. The specific operation steps are as follows: Randomly divide the dataset into 10 equal parts in each fold cross validation, and take 1 part as the test set and the remaining 9 parts as the training set. Repeat the above process 10 times, and take the average of the 10 results as the evaluation index.

4.4. Comparison of Mark Distribution Algorithms.
Table 5 summarizes the experimental results of six labeled distribution learning algorithms and CSLDSVR on five different datasets, and the experimental results are recorded in the form of mean ± standard deviation. Among them, the bold is the best value of each indicator in different methods on the current dataset. Clearly, in comparison with the label distribution learning algorithm, CSLDSVR has shown excellent results in most cases, and it is more obvious on the UM, UCLA, and KKI datasets. Among the indicators of the

Table 5: Comparative performance evaluation of CSLDSVR and LDL algorithms.

Evaluation metrics	Algorithm	NYU	UM	Leuven	UCLA	KKI
Chebyshev↓	AA-BP	0.223 7 ± 0.035 6	0.218 4 ± 0.045 8	0.248 0 ± 0.044 6	0.250 6 ± 0.053 5	0.254 7 ± 0.052 9
	AA-*K*NN	0.144 1 ± 0.011 6	0.154 0 ± 0.021 1	0.157 9 ± 0.026 5	0.142 6 ± 0.031 3	0.157 2 ± 0.029 5
	LDSVR	0.150 1 ± 0.024 3	0.140 0 ± 0.012 8	0.162 9 ± 0.034 4	0.169 4 ± 0.053 4	0.160 2 ± 0.057 0
	SA-IIS	0.147 8 ± 0.011 8	0.153 5 ± 0.023 7	0.174 8 ± 0.021 4	0.145 8 ± 0.032 9	0.162 7 ± 0.049 5
	PT-BAYES	0.381 8 ± 0.111 9	0.205 7 ± 0.009 5	0.206 9 ± 0.007 8	0.213 5 ± 0.009 9	0.215 4 ± 0.008 1
	PT-SVM	0.200 5 ± 0.041 2	0.188 5 ± 0.042 3	0.183 1 ± 0.040 1	0.195 8 ± 0.033 0	0.182 2 ± 0.058 9
	CSLDSVR	**0.141 3 ± 0.016 2**	**0.135 2 ± 0.023 6**	**0.140 2 ± 0.024 4**	**0.138 6 ± 0.038 4**	**0.126 7 ± 0.034 9**
Cosine↑	AA-BP	0.873 1 ± 0.034 4	0.881 8 ± 0.035 6	0.862 2 ± 0.049 8	0.839 9 ± 0.057 8	0.843 7 ± 0.058 6
	AA-*K*NN	0.935 4 ± 0.009 6	0.928 6 ± 0.017 3	0.927 4 ± 0.020 8	0.929 7 ± 0.022 4	0.913 0 ± 0.024 4
	LDSVR	0.937 7 ± 0.019 1	0.944 8 ± 0.013 3	**0.932 5 ± 0.029 2**	0.928 5 ± 0.052 0	0.932 6 ± 0.047 4
	SA-IIS	**0.940 7 ± 0.009 3**	0.934 4 ± 0.016 7	0.920 5 ± 0.016 0	0.939 5 ± 0.020 3	0.924 6 ± 0.042 5
	PT-BAYES	0.798 5 ± 0.071 3	0.915 6 ± 0.006 2	0.915 1 ± 0.005 3	0.910 4 ± 0.006 9	0.909 2 ± 0.005 7
	PT-SVM	0.898 7 ± 0.038 5	0.904 3 ± 0.042 8	0.914 5 ± 0.030 9	0.897 4 ± 0.036 5	0.906 8 ± 0.045 8
	CSLDSVR	0.940 5 ± 0.012 1	**0.947 3 ± 0.018 3**	0.923 4 ± 0.025 5	**0.942 8 ± 0.036 8**	**0.936 3 ± 0.029 4**
Clark↓	AA-BP	0.468 1 ± 0.064 8	0.461 3 ± 0.099 0	0.517 0 ± 0.083 8	0.537 1 ± 0.110 1	0.542 7 ± 0.104 6
	AA-*K*NN	0.263 1 ± 0.020 3	0.282 2 ± 0.036 7	0.287 3 ± 0.047 3	0.261 3 ± 0.053 5	0.283 2 ± 0.053 9
	LDSVR	0.272 9 ± 0.036 4	0.255 7 ± 0.021 8	0.287 2 ± 0.062 6	0.295 6 ± 0.092 0	0.281 9 ± 0.100 8
	SA-IIS	0.266 3 ± 0.019 1	0.278 8 ± 0.039 7	0.311 3 ± 0.033 6	0.262 3 ± 0.055 5	0.293 9 ± 0.088 0
	PT-BAYES	0.893 6 ± 0.359 8	0.352 0 ± 0.014 5	0.352 3 ± 0.012 7	0.363 6 ± 0.016 2	0.366 3 ± 0.013 3
	PT-SVM	0.358 0 ± 0.070 2	0.348 1 ± 0.075 8	0.325 3 ± 0.065 5	0.350 5 ± 0.056 1	0.328 7 ± 0.098 1
	CSLDSVR	0.261 6 ± 0.032 1	0.246 3 ± 0.037 6	**0.253 9 ± 0.041 8**	**0.248 4 ± 0.062 6**	**0.233 4 ± 0.061 8**
Intersection↑	AA-BP	0.776 3 ± 0.035 6	0.781 6 ± 0.045 8	0.752 0 ± 0.044 6	0.749 4 ± 0.053 5	0.745 3 ± 0.052 9
	AA-*K*NN	0.855 9 ± 0.011 6	0.846 0 ± 0.021 1	0.842 1 ± 0.026 5	0.857 4 ± 0.031 3	0.842 8 ± 0.029 5
	LDSVR	0.849 9 ± 0.024 3	0.860 0 ± 0.012 8	0.837 1 ± 0.034 4	0.830 6 ± 0.053 4	0.839 8 ± 0.057 0
	SA-IIS	0.852 2 ± 0.011 8	0.846 5 ± 0.023 7	0.825 2 ± 0.021 4	0.854 2 ± 0.032 9	0.837 3 ± 0.049 5
	PT-BAYES	0.618 2 ± 0.111 9	0.794 3 ± 0.009 5	0.793 1 ± 0.007 8	0.786 5 ± 0.009 9	0.784 6 ± 0.008 1
	PT-SVM	0.799 5 ± 0.041 2	0.811 5 ± 0.042 3	0.816 9 ± 0.040 1	0.804 2 ± 0.033 0	0.817 8 ± 0.058 9
	CSLDSVR	0.858 7 ± 0.041 5	0.864 8 ± 0.023 6	0.859 8 ± 0.024 4	0.861 4 ± 0.038 4	0.873 3 ± 0.034 9
KL↑	AA-BP	0.166 7 ± 0.042 9	0.161 2 ± 0.051 7	0.192 0 ± 0.069 3	0.222 2 ± 0.089 8	0.227 9 ± 0.076 4
	AA-*K*NN	0.068 5 ± 0.010 1	0.076 0 ± 0.018 4	0.076 6 ± 0.022 1	0.074 6 ± 0.023 2	0.093 2 ± 0.026 6
	LDSVR	0.066 5 ± 0.019 9	0.059 3 ± 0.014 6	0.070 3 ± 0.032 3	0.074 9 ± 0.062 5	0.071 1 ± 0.049 8
	SA-IIS	0.063 9 ± 0.009 3	0.069 8 ± 0.017 8	0.083 7 ± 0.016 6	0.063 9 ± 0.021 0	0.080 0 ± 0.044 1
	PT-BAYES	0.492 9 ± 0.251 0	0.087 9 ± 0.006 7	0.088 0 ± 0.006 1	0.093 5 ± 0.008 0	0.094 8 ± 0.006 6
	PT-SVM	0.108 1 ± 0.041 2	0.105 5 ± 0.047 6	0.090 6 ± 0.032 9	0.110 4 ± 0.040 5	0.100 3 ± 0.048 1
	CSLDSVR	0.060 3 ± 0.041 5	0.056 7 ± 0.019 5	0.069 9 ± 0.024 0	0.060 1 ± 0.046 1	0.068 2 ± 0.030 1

labeled distribution, KL divergence is an indicator describing the difference between the two distributions, and the LDL algorithm used as a comparison uses KL divergence as the objective function. The KL divergence of the prediction result of CSLDSVR can be minimized. It shows that the label distribution predicted by the new algorithm is the closest to the real data distribution on the whole, which is better than the comparison algorithm.

Figure 2 summarizes the results of CSLDSVR and the marker distribution algorithm multiclass metrics precision and mAP; from the two most important multiclass metrics, CSLDSVR performs better. Some algorithms have a high accuracy rate but a low macro average because these algorithms do not consider the class imbalance problem, and the model classification is biased towards the majority class. CSLDSVR uses the kernel trick to solve the problem in a more discriminative feature space, and CSLDSVR considers the size of each class, which effectively solves the problem caused by class imbalance.

To verify the performance improvement of the cost-sensitive mechanism, the algorithm in this paper is compared with the
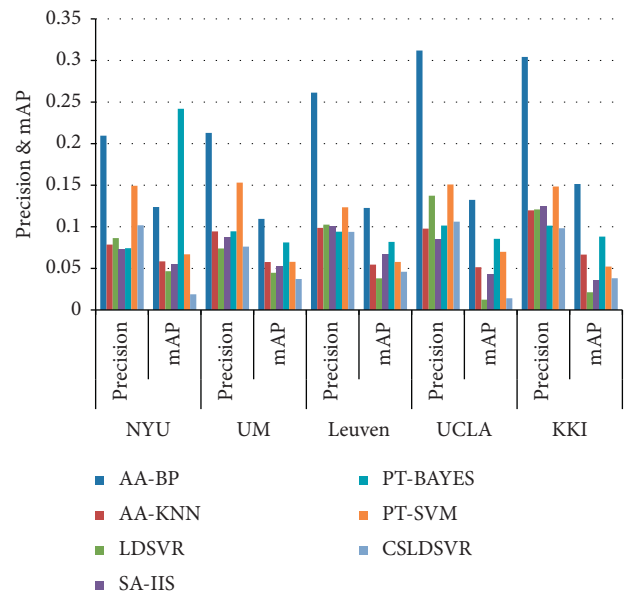
Figure 2: Comparative performance evaluation of CSLDSVR and label distribution algorithm.

TABLE 6: Performance evaluation for multiclassification algorithms.

Dataset	Decision tree		KNN		CSLDSVR	
	Precision	mAP	Precision	mAP	Precision	mAP
NYU	0.548 8 ± 0.1423	0.409 3 ± 0.0703	0.614 4 ± 0.1525	0.364 7 ± 0.0527	**0.655 4 ± 0.0571**	**0.451 7 ± 0.0398**
UM	0.576 7 ± 0.1325	0.385 9 ± 0.0872	0.528 5 ± 0.1214	0.374 0 ± 0.0861	**0.701 4 ± 0.0708**	**0.497 1 ± 0.1250**
Leuven	0.617 1 ± 0.2261	0.424 2 ± 0.2086	0.608 5 ± 0	0.333 3 ± 0	**0.617 6 ± 0.0725**	**0.448 2 ± 0.0861**
UCLA	0.605 2 ± 0.1833	0.442 0 ± 0.2086	0.654 3 ± 0	0.333 3 ± 0	**0.665 2 ± 0.1504**	**0.443 4 ± 0.1659**
KKI	0.559 8 ± 0.2567	0.395 4 ± 0.2941	0.646 5 ± 0	0.333 3 ± 0	**0.687 5 ± 0.1237**	**0.447 6 ± 0.1016**
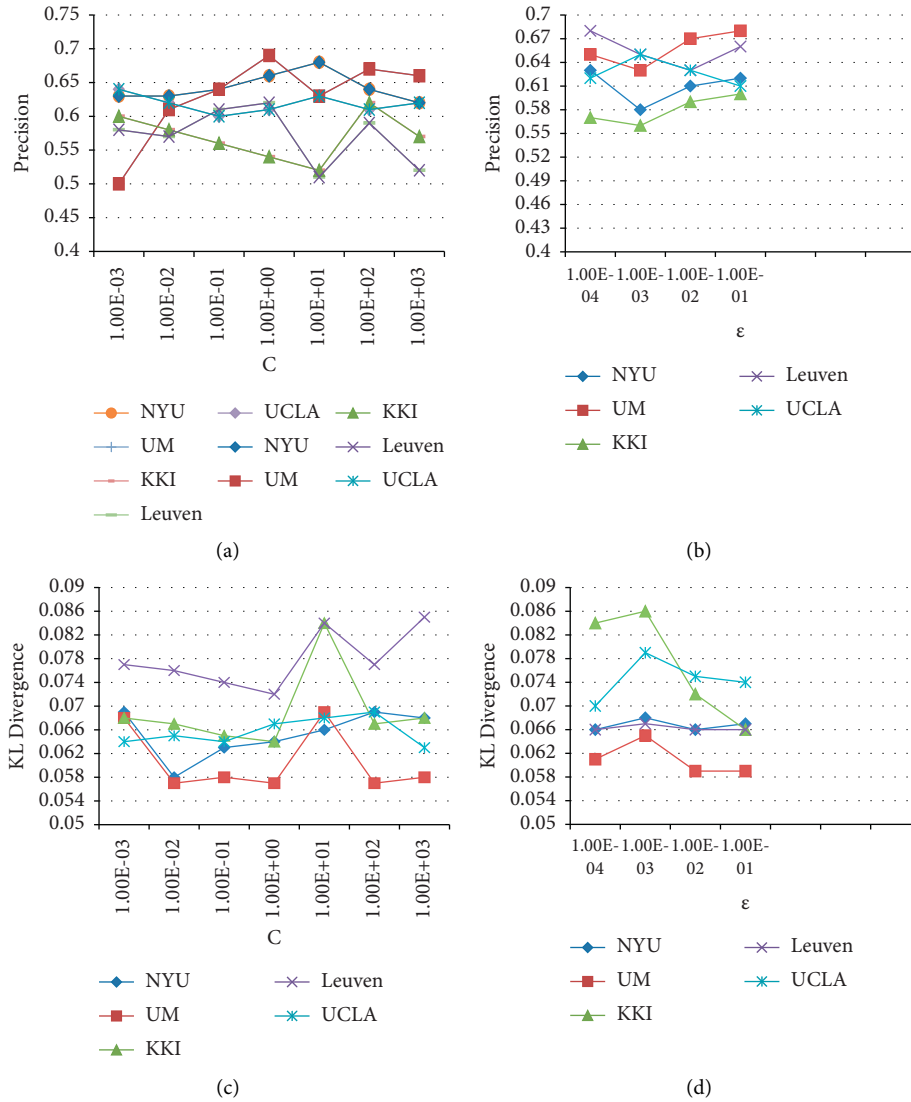
FIGURE 3: Changes of evaluation indicators. (a) Impact of C on precision, (b) impact of ε on precision, (c) impact of C on KL, and (d) impact of ε on KL.

LDSVR without the cost-sensitive mechanism. As shown in Table 5, in most cases, the learning effect of the algorithm CSLDSVR in this paper is better; in addition, the standard deviation of the results is basically maintained at a low level, that is, the stability of the algorithm is improved. However, LDSVR does not introduce a cost-sensitive mechanism, and the standard deviation of the results obtained by the algorithm is large and

fluctuating. For example, the standard deviation of the Canberra indicators in UCLA and KKI exceeds 0.1.

4.5. *Multiclass Comparison Experiment.* Table 6 shows the comparison results of precision and mAP metrics of CSLDSVR and two classical multiclassification algorithms,

decision tree and KNN, on five datasets. Among them, the bold is the best value of the corresponding indicator in different methods on the current dataset. Observing the experimental results of the KNN method, it can be found that the mAP of the KNN method appears 0.333 3 times, this is because KNN is too biased towards the majority class, and there is an extreme case of classifying all samples into the majority class. In the case of high-dimensional imbalance of autism neuroimaging data, traditional multiclassification algorithms are prone to fall into the dimensional trap or bias towards the majority class. The algorithm CSLDSVR in this paper solves the above problems by using kernel skills and cost-sensitive mechanisms and achieves better results. Good classification model. The cost-sensitive mechanism reduces the overall misclassification cost by increasing the misclassification cost of the minority class and reducing the misclassification cost of the majority class and makes the model avoid leaning towards the majority class. In other words, the cost-sensitive mechanism is based on the original standard cost loss function, adding some constraints and weight constraints, so that the final model is biased towards another minority class that is more concerned in practical applications. This paper achieves the purpose of different misjudgment costs for different categories by introducing $1Nj$. In theory, this can avoid the tendency of the algorithm model to the majority class and improve the prediction accuracy for the minority class [37]. In the experiment, the experimental results in Table 6 also verify this theory, and in most cases, the stability of the algorithm has also been improved, and the standard deviation of the experimental results is small.

4.6. Effect of Parameters. In this section, we study the effect of parameter changes on the performance of the algorithm CSLDSVR. Figure 3 shows the changes of the evaluation indicators precision and KL divergence when the parameters C and ε take different values on five different datasets. Comparing and studying two graphs of the same parameter and different indicators, such as Figures 3(a) and 3(c), it can be found that the curve trend of the same dataset is basically opposite, and the point where precision takes the maximum value is generally the same as the KL divergence is the minimum value, which also corresponds to the previous analysis of KL divergence, indicating that when the KL divergence is small, the label distributions of the two are more similar, and the classification results are more accurate.

It is found that for different datasets, the parameter values for obtaining the optimal solution are not the same, which also shows that in the diagnosis of autism, the data distribution of different data centers is different, and the parameters for building the model should also be different. Moreover, it is found that for a dataset with fewer samples, the result is more sensitive to the change of the parameters, such as for the KKI dataset with only 48 samples, the fluctuation is the largest when the parameter value changes.

It can be seen that the parameters of the CSLDSVR algorithm should be based on the characteristics of the dataset, and the corresponding parameter values should be

set to build a model. If the parameter settings are reasonable, CSLDSVR can overcome the high dimensionality and category imbalance of the autism dataset. Thus, the whole section contains the strategy of how ASD detection is done by evaluating several strategies such as SVR and LDL by considering certain parameters. Therefore, the label distribution learning approach overcomes high-dimensional feature classification issues by mapping samples to the feature space and then diagnosing multiclass ASD. This technique outperforms previous methods in terms of classification performance and accuracy, as well as resolving the issue of unbalanced data in ASD diagnosis.

5. Conclusion

This research presents a cost-sensitive marker distribution to enable an ASD-aided diagnostic approach for vector regression based on functional connectivity characteristics collected from rs-fMRI. Since ASD patients' brain function differs from that of healthy persons,so rs-fMRI is a useful method for capturing brain activity. In this study, researchers have introduced the label distribution learning that solves the label noise problem in multiclassification ASD diagnosis. Furthermore, the new technique have been implemented, which provides class balancing and in addition balances the effect of the majority and minority classes on the objective function using the labeled distribution support vector regression method. The new method employed in this study effectively solves the imbalanced data problem in ASD diagnosis by overcoming the imbalance of the influence of the majority and minority classes on the results obtained in the paper. Besides, it presents a cost-sensitive approach to correct sample imbalance and uses a support vector regression (SVR)-based method to remove label noise. The label distribution learning approach overcomes high-dimensional feature classification issues by mapping samples to the feature space and then diagnosing multiclass ASD. The overall result obtained in this technique outperforms previous methods in terms of classification performance and accuracy, as well as resolves the issue of unbalanced data in ASD diagnosis. However, the improved model is still biased towards the majority class to some extent, and the imbalanced data problem should be improved further as a future study. Researchers can further try to improve the data sampling method or use the synthetic minority sample method, etc. as future prespective. However, relatively high-level distances must also be introduced, which necessitates more prior knowledge. Since prior knowledge is no longer used, the Euclidean distance is used instead. Other advanced distances have their set of benefits that will be refined in future research.

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] M. Greicius, "Resting-state functional connectivity in neuropsychiatric disorders," *Current Opinion in Neurology*, vol. 21, no. 4, pp. 424–430, 2008.

[2] L. Cerliani, M. Mennes, R. M. Thomas, A. Di Martino, M. Thioux, and C. Keysers, "Increased functional connectivity between subcortical and cortical resting-state networks in autism spectrum disorder," *JAMA Psychiatry*, vol. 72, no. 8, p. 767, 2015.

[3] 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, 16 pages, 2021.

[4] S. Broder Fingert, A. Carter, K. Pierce et al., "Implementing systems-based innovations to improve access to early screening, diagnosis, and treatment services for children with autism spectrum disorder: an Autism Spectrum Disorder Pediatric, Early Detection, Engagement, and Services network study," *Autism*, vol. 23, no. 3, pp. 653–664, 2018.

[5] G. Kaur Saini, H. Chouhan, S. Kori et al., "Recognition of human sentiment from image using machine learning," *Annals of the Romanian Society for Cell Biology*, vol. 25, 2021.

[6] P. Lichtenstein, E. Carlström, M. Råstam, C. Gillberg, and H. Anckarsäter, "The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood," *American Journal of Psychiatry*, vol. 167, no. Issue 11, pp. 1357–1363, 2010.

[7] A. Ronald and R. A. Hoekstra, "Autism spectrum disorders and autistic traits: a decade of new twin studies," in *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* vol. 156, no. Issue 3, pp. 255–274, Wiley, 2011.

[8] J. Chen, L. Chen, and M. Shabaz, "Image fusion algorithm at pixel level based on edge detection," *Journal of Healthcare Engineering*, vol. 2021, Article ID 5760660, 10 pages, 2021.

[9] T. Gaugler, L. Klei, S. J. Sanders et al., "Most genetic risk forautism resides with common variation," *Nature Genetics*, vol. 46, no. 8, pp. 881–885, 2014.

[10] A. O. Caglayan, "Genetic causes of syndromic and non-syndromic autism," in *Developmental Medicine and Child Neurology,* vol. 52, no. Issue 2, pp. 130–138, Wiley, 2010.

[11] I. Iossifov, B. J. O'Roak, S. J. Sanders et al., "The contribution of denovo coding mutations to autism spectrum disorder," *Nature*, vol. 515, no. 7526, pp. 216–221, 2014.

[12] Q. Yao, M. Shabaz, T. K. Lohani, M. Wasim Bhatt, G. S. Panesar, and R. K. Singh, "3D modelling and visualization for vision-based vibration signal processing and measurement," *Journal of Intelligent Systems*, vol. 30, no. Issue 1, pp. 541–553, 2021.

[13] A. Ronald, F. Happé, P. Bolton et al., "Genetic heterogeneity between the three components of the autism spectrum: a twin study," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 45, no. Issue 6, pp. 691–699, 2006.

[14] R. H. Wozniak, N. B. Leezenbaum, J. B. Northrup, K. L. West, and J. M. Iverson, "The development of autism spectrum disorders: variability and causal complexity," in *Wiley Interdisciplinary Reviews: Cognitive Science,* vol. 8, Article ID e1426, 2016.

[15] A. Kumar, A. mehbodniya, J. L. Webber et al., "Optimal cluster head selection for energy efficient wireless sensor network using hybrid competitive swarm optimization and harmony search algorithm," *Sustainable Energy Technologies and Assessments*, vol. 52, no. 102243, pp. 102243–102245, 2022.

[16] J. T. Elison, S. J. Paterson, J. J. Wolff et al., "White matter microstructure and atypical visual orienting in 7-month-olds at risk for autism," *American Journal of Psychiatry*, vol. 170, no. 8, pp. 899–908, 2013.

[17] G. Esposito, P. Venuti, S. Maestro, and F. Muratori, "An exploration of symmetry in early autism spectrum disorders: analysis of lying," *Brain & Development*, vol. 31, no. Issue 2, pp. 131–138, 2009.

[18] E. J. H. Jones, T. Gliga, R. Bedford, T. Charman, and M. H. Johnson, "Developmental pathways to autism: a review of prospective studies of infants at risk," *Neuroscience & Biobehavioral Reviews*, vol. 39, pp. 1–33, 2014.

[19] M. A. Gernsbacher, J. L. Stevenson, S. Khandakar, and H. H. Goldsmith, "Why does joint attention look atypical in autism?" in *Child Development Perspectives,* vol. 2, no. Issue 1, pp. 38–45, Wiley, 2008.

[20] D. Fein, M. Barton, I.-M. Eigsti et al., "Optimal outcome in individuals with a history of autism," *Journal of Child Psychology and Psychiatry*, vol. 54, no. Issue 2, pp. 195–205, 2013.

[21] E. Courchesne and K. Pierce, "Brain overgrowth in autism during a critical time in development: implications for frontal pyramidal neuron and interneuron development and connectivity," *International Journal of Developmental Neuroscience*, vol. 23, pp. 153–170, 2005.

[22] C. Frith, "Is autism a disconnection disorder?" *The Lancet Neurology*, vol. 3, no. Issue 10, p. 577, 2004.

[23] M. A. Just, V. L. Cherkassky, T. A. Keller, R. K. Kana, and N. J. Minshew, "Functional and anatomical cortical under-connectivity in autism: evidence from an fMRI study of an executive function task and corpus callosum morphometry," *Cerebral Cortex*, vol. 17, no. Issue 4, pp. 951–961, 2006.

[24] B. C. S. Mindblindness, "An Essay on AutismandTheoryofMind. Cambridge, MA: MITPress; 1995.21. Ozonoff S, Pennington BF, Rogers SJ. Executive func-tion deficits in high-functioning autistic individuals:relationship to theory of mind," *J Child Psychol Psychi-atry1991*, vol. 32, pp. 1081–1105.22.

[25] F. Happé, "Autism: cognitive deficit or cognitive style?" *Trends in Cognitive Sciences*, vol. 3, no. Issue 6, pp. 216–222, 1999.

[26] V. Vishal, A. Singh, Y. B. Jinila, S. P. Shyry, and J. Jabez, "A comparative analysis of prediction of autism spectrum disorder (ASD) using machine learning," in *Proceedings of the 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, IEEE, Tirunelveli, India, April 2022, https://doi.org/10.1109/icoei53556.2022.9777240.

[27] W. Yin, L. Li, and F.-X. Wu, "A semi-supervised autoencoder for autism disease diagnosis," in *Neurocomputing,* vol. 483, pp. 140–147, Elsevier BV, 2022.

[28] M. A. Haq, "Probabilistic methods and neural networks in structural engineering," *International Journal of Advanced Manufacturing Technology*, vol. 120, no. 11, pp. 1–9, 2022.

[29] X. Geng, C. Yin, and Z.-H. Zhou, "Facial age estimation by learning from label distributions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. Issue 10, pp. 2401–2412, 2013.

[30] A. Gupta, L. Kapoor, and M. Wattal, "C2C (cloud-to-cloud): an ecosystem of cloud service providers for dynamic resource provisioning," in *Advances in Computing and Communications*, pp. 501–510, Springer, Berlin Heidelberg, 2011.

[31] L. Kapoor, S. Bawa, and A. Gupta, "Peer clouds: a P2P-based resource discovery mechanism for the Intercloud,"

International Journal of Next-Generation Computing, vol. 12, pp. 153–164, 2015.

[32] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang, "Head Pose Estimation: Classification or Regression? in 2008 19th International Conference on Pattern Recognition," in *Proceedings of the 2008 19th International Conference on Pattern Recognition (ICPR)*, January 2008.

[33] S. Gurbuz, E. Oztop, and N. Inoue, "Model free head pose estimation using stereovision," *Pattern Recognition*, vol. 45, no. Issue 1, pp. 33–42, 2012.

[34] M. A. Haj, J. Gonzalez, and L. S. Davis, "On partial least squares in head pose estimation: how to simultaneously deal with misalignment," in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, June 2012.

[35] T. Thakur, I. Batra, M. Luthra et al., "Gene expression-assisted cancer prediction techniques," *Journal of Healthcare Engineering*, vol. 2021, p. 1, Article ID 4242646, 2021.

[36] A. Gupta and P. Prabhat, "Novel approaches in network fault management," *INTERNATIONAL JOURNAL OF NEXT-GENERATION COMPUTING*, vol. 8, no. 2, 2017.

[37] S. Lawrence, I. Burns, A. Back, A. Tsoi, and C. Giles, "Neural network classification and prior class probabilities," *Neural Networks: Tricks of the Trade*, vol. 1524, pp. 299–313, 1998.

Research Article

Development of Machine Learning and Medical Enabled Multimodal for Segmentation and Classification of Brain Tumor Using MRI Images

L. Anand ⬥,[1] Kantilal Pitambar Rane,[2] Laxmi A. Bewoor,[3] Jyoti L. Bangare,[4] Jyoti Surve,[5] Mutkule Prasad Raghunath,[6] K. Sakthidasan Sankaran ⬥,[7] and Bernard Osei ⬥[8]

[1]*Department of Networking and Communications, SRM Institute of Science and Technology, Chennai, India*
[2]*Department of Electronics and Communication, Koneru Lakshmaiah Education Foundation (Deemed to Be University), Andhra Pradesh, Vaddeswaram, India*
[3]*Department of Computer Engineering, Vishwakarma Institute of Information Technology, Savitribai Phule Pune University, Pune, India*
[4]*Department of Computer Engineering, MKSSS's Cummins College of Engineering for Women, Savitribai Phule Pune University, Pune, India*
[5]*Department of Information Technology, International Institute of Information Technology, Hinjewadi, Savitribai Phule Pune University, Pune, India*
[6]*Department of Information Technology, SRES's Sanjivani College of Engineering, Kopargaon 423603, Maharashtra, India*
[7]*Department of ECE, Hindustan Institute of Technology and Science, Chennai, India*
[8]*Kwame Nkrumah University of Science and Technology, Kumasi, Ghana*

Correspondence should be addressed to Bernard Osei; bosei26@st.knust.edu.gh

The improper and excessive growth of brain cells may lead to the formation of a brain tumor. Brain tumors are the major cause of death from cancer. As a direct consequence of this, it is becoming more challenging to identify a treatment that is effective for a specific kind of brain tumor. The brain may be imaged in three dimensions using a standard MRI scan. Its primary function is to examine, identify, diagnose, and classify a variety of neurological conditions. Radiation therapy is employed in the treatment of tumors, and MRI segmentation is used to guide treatment. Because of this, we are able to assess whether or not a piece that was spotted by an MRI is a tumor. Using MRI scans, this study proposes a machine learning and medically assisted multimodal approach to segmenting and classifying brain tumors. MRI pictures contain noise. The geometric mean filter is utilized during picture preprocessing to facilitate the removal of noise. Fuzzy c-means algorithms are responsible for segmenting an image into smaller parts. The identification of a region of interest is facilitated by segmentation. The GLCM Grey-level co-occurrence matrix is utilized in order to carry out the process of dimension reduction. The GLCM algorithm is used to extract features from photographs. The photos are then categorized using various machine learning methods, including SVM, RBF, ANN, and AdaBoost. The performance of the SVM RBF algorithm is superior when it comes to the classification and detection of brain tumors.

1. Introduction

The improper and excessive growth of brain cells may lead to the formation of a brain tumor. Brain tumors are the major cause of death from cancer in those under the age of 19, accounting for 24 percent of all deaths caused by cancer.

There are around 120 very different types of brain tumors. As a direct consequence of this, it is becoming more challenging to identify a treatment that is effective for a specific kind of brain tumor. For the purposes of categorization, brain tumors may be broken down into two primary subtypes: benign and malignant. Benign brain tumors are less

dangerous than their malignant counterparts. When discussing the many kinds of cancer, the term "benign" refers to a tumor that does not metastasize, or spread, to other regions of the body. This kind of cancer is less likely to be fatal. Excision surgery is the standard treatment for benign tumors, which can be cured in most cases. Tumors that are cancerous have a high risk of spreading to other areas of the body and are notoriously difficult to remove [1, 2].

However, if the patient has any kind of brain tumor, whether it be malignant or benign, the patient is in danger and may even pass away. This is due to the fact that benign brain tumors may not invade the surrounding tissue, but as they get bigger, they place additional pressure on neighboring brain cells that are essential for normal brain function. This can cause brain damage. Brain tumors may originate in the tissue of the brain itself, or they might be brought there by cancer cells that have traveled to the brain from another part of the body. The process by which tumor cells from one source go to another organ and infiltrate the tissue of that new place is referred to as "metastasis," and this is the meaning of the word "metastasis" [3, 4].

Using a regular MRI scan, a three-dimensional image of the human brain is possible to obtain. The major purpose of this organization is to investigate, recognize, diagnose, and categorize a wide range of neurological diseases. Radiation therapy is typically used in the treatment of cancers, and MRI segmentation is frequently utilized to direct radiation therapy. As a result of this, we are in the position to determine whether or not a fragment that was detected by an MRI is a tumor. These treatments do not involve the use of ionizing radiation and are instead used for diagnosis, the identification of disease stages, and subsequent monitoring. The most common setting in which these treatments are used is within hospitals. Magnetic resonance imaging (MRI), one of the most beneficial modalities that are presently available, is used to diagnose more advanced stages of brain tumors [5].

When looking for potentially hazardous areas in medical images, segmentation is an essential step that must be taken. There is potential for tremendous use and precision in the use of automated identification for the early diagnosis of brain tumors by MRI. The MRI of the brain tumor is shown in Figure 1.

The field of artificial intelligence encompasses a wide range of subfields, including machine learning and computational theory, searching and probability, as well as neuroscience [6]. To begin, it is fairly excellent at obtaining homogenous data from a big number of data, and it is able to learn from any form of data, whether it be numeric, visual, or video data. In addition, it can learn from any amount of data. Getting familiar with a dataset might provide you an advantage in the research and data extraction work you are doing. At the moment, those working in the field of medicine are having trouble diagnosing human diseases at an early stage so effective medicines may be created to assist patients in living longer lives. The identification of diabetic retinopathy, breast cancer, and lung cancer and the detection and diagnosis of brain tumors using MRI are some of the medical applications that make use of prescreening. It is
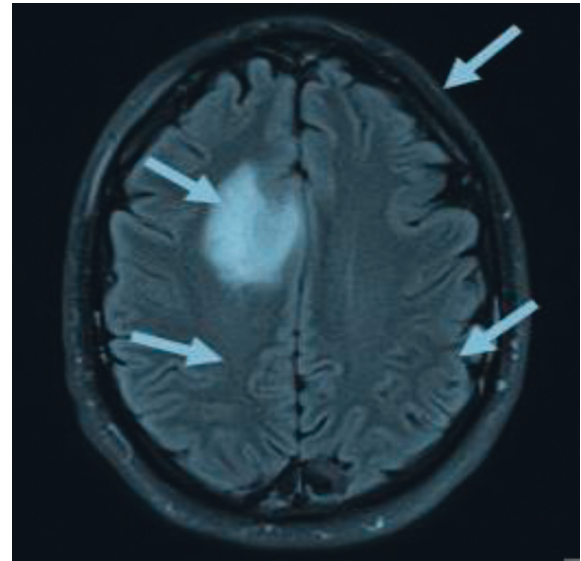
FIGURE 1: MRI image of brain tumor.

very necessary to make use of image processing methods that facilitate automatic learning and the extraction of information. AI strategies have the potential to assist addressing a larger variety of complicated problems in a more timely and efficient manner [7, 8].

The literature review section contains a survey of existing techniques for brain tumor detection. The methodology section describes how to use MRI scans to segment and classify brain tumors using machine learning and medically enabled multimodal. Noise is seen in MRI pictures. The geometric mean filter is used in picture preprocessing to eliminate noise. FCM (fuzzy c-means) algorithms split images into parts. The identification of an area of interest is aided by segmentation. The GLCM Grey-level co-occurrence matrix is used to reduce the dimensions. GLCM is a program that extracts characteristics from photos. Finally, machine learning methods such as SVM RBF, ANN, and AdaBoost are used to classify pictures. The result section presents results achieved by various machine learning and feature extraction techniques. The conclusion section contains major contributions of the article to brain tumor detection.

2. Literature Survey

This section presents a literature survey of image preprocessing, feature extraction, segmentation, and classification techniques in the context of brain tumor detection.

2.1. Literature Survey of Image Preprocessing Techniques. The process of denoising MRI images was the focus of research that Agrawal and Sahu conducted [9]. Denoising an MRI picture was accomplished by the author via the use of the discrete wavelet transform (DWT). Wavelet analysis may be used to address discontinuities in higher derivatives, breakdown points, and other self-similarities that are overlooked by standard signal analysis approaches. It is able

to denoise and compress a signal without compromising the signal's quality. The discrete wavelet transform not only creates an image that is free of unnecessary redundancies, but also provides additional spectral and spatial information. When compared to the information offered by the Gaussian or Laplacian pyramid, the information that it delivers is richer and more detailed. In the DWT method, low and high pass filters are responsible for the creation of approximation and detail coefficients. DWT may also be used to do the technique in reverse, which will restore the original signal in its entirety with no degradation.

An article on methods for improving medical image processing was authored by Bedi and Rati Khandelwal [10], and it was titled "Medical Image Processing Improvement Approaches." The likes of medical pictures, aerial photography, and other sorts of imaging all suffer from issues with poor contrast and noisy signal. Increasing an image's contrast and sharpness and minimizing blurriness and making an image more obvious to the naked eye are some of the methods that may be used to improve picture quality. The spatial and frequency domain techniques are the two categories of strategies that may be used to enhance images. The quality of the image was enhanced by the use of spatial domain techniques including negative transforms, power-log transformations, piecewise slicing, and histograms. In order to produce smooth pictures in the frequency domain, Fourier transformations of an existing picture are used. These changes reduce the intensity of a predetermined range of frequency components.

Mathen and George are the ones who created a method for improving and denoising medical images [11]. The technique that was proposed consisted of three stages: preprocessing, enhancing the contrast, and denoising the image. The authors used the median filter in order to lessen the amount of noise and increase the level of clarity towards the margins. In the second phase, the histogram equalization method is used in order to increase the contrast of the image. To ensure that the grayscale values are distributed consistently over the picture, a preprocessing technique called as histogram equalization is used. The contrast of a picture that seems natural is improved by doing this. The histogram of a picture should be stretched in order to achieve the purpose of histogram equalization, which is to make the histogram as flat as possible. In an image of the brain taken with an MRI, it improves the contrast between the grey matter, the white matter, and the cerebrospinal fluid (CSF). At the very end of the procedure, the images were denoised by using a nonlocal means filter. It was able to reduce noise while maintaining the quality of the small details. The developer of the nonlocal means filter had the opinion that the image had a high degree of likeness to itself in the context of this situation.

2.2. Literature Survey of Image Segmentation Techniques. Dina et al. [12] demonstrated how a modified image segmentation approach might be used to analyze MR images for the purpose of locating brain tumors. Their learning vector quantization (LVQ)-based modified probabilistic neural network (PNN) model surpassed all others in training,

classification with a hundred percent accuracy, and a decrease in compute time of seventy-nine percent in the processing of images and data. It has been shown and researched how linear and Gaussian filters work, in addition to techniques of augmentation and smoothing. The Canny edge detection technique was also used in the research on edge detection at various points.

Leela and Veenakumari [13] developed a computer-aided detection method for aberrant tissue growth with the purpose of improving medical diagnosis. This method involves the processing of MR images with the highest possible accuracy and speed. Additionally, they have improved the image quality by cutting down on the amount of noise that was there. The images of brain tumors, as well as morphological processing and segmentation procedures, have all been explored for their qualities of discontinuous or comparable intensity values.

In their 2014 study, Rohini Paul et al. [14] suggested using the K-means clustering approach to separate the data obtained from brain MRI scans. In order to prevent the formation of clustered areas during the segmentation of brain MRI images for the purpose of diagnosing tumors, morphological filtering is an absolute need.

This is the first time that Hamoud et al. have offered [15] a complete evaluation of the technique and methods that were used. Motifs and edges wrap up with a perceptive summary that paves the way for further study on brain picture segmentation, thresholding, and noise reduction, as well as cancer diagnostics.

Using photographs of brain tumors, Siva Sankari et al. developed a technique for segmenting the images and extracting information from them [16]. The K-means clustering method was used so that this could be accomplished. Reddy et al. [17] have devised a method for estimating the volume of a tumor by making use of the data from multi-parametric MR images, in addition to coming up with the novel concept of a confidence surface to assist in segmentation. In order to train the classifier and the segmentation approach, contrast-weighted image information and texture information were both used.

Using quantitative methodologies for automated feature identification based on the identity of each pixel in the picture, Megha et al. [18] demonstrated how to improve visual discrimination between scene features and replace visual analysis of image data with quantitative methods. Additionally, they demonstrated how to increase visual discrimination between scene features. Extraction of data via the use of digital image processing has also been considered.

Researchers Hemlatha et al. [19] discovered that it was possible to automatically establish the location of a tumor in the brain as well as its size by using MRI technology. They have shown how to do this via the use of digital picture processing.

2.3. Literature Survey of Feature Extraction Techniques. Mohanaiah et al. proposed a technique for image processing that retrieves textural information from a picture [20]. The Grey-level co-occurrence matrix (GLCM) approach was

used in order to extract textural properties. Image analysis would not be complete without the use of texture analysis in the picture-sorting process. The GLCM algorithm may be used to extract textural properties from grayscale photographs. Ailments may be categorized as either normal or abnormal by applying these factors in the software used in medicine. This statistical method allows for the extraction of features. The author used this matrix technique in order to evaluate the properties of image movements. In order to evaluate properties such as angular second moment (energy), correlation, and entropy, an MRI image was used. This approach is more efficient in terms of time consumption when compared to the discrete wavelet transform.

In an article, Ping Tian and colleagues [21] proposed a method for extracting color image attributes. This method was included in the work. In RGB color photographs, a pixel's color is considered a property, but in greyscale photos, the texture is considered a property of an entire group of pixels. Images and the domains they belong to are used in order to obtain the statistical or pixel structure of a texture. This study looks at two different kinds of characteristics: local and global. Both of these are analyzed in this research. The LBP, Gabor wavelet, and histogram methods are used in this study to extract the characteristics of color images. The histogram method may be used in order to extract global features from an image. On the other hand, the Gabor wavelet and the Sobel shape detector can be utilized in order to extract local features from an image. When making the image portrayal, these considerations were taken into account; thus, the result is accurate.

An algorithm for the identification of brain tumors was suggested by Kourosh Jafari [22]. This approach makes use of high-resolution (HR) images with various degrees of contrast. The majority of the time, these low-contrast photographs are upsampled with high-contrast images to produce higher resolution versions of the original photos. The algorithm that is being presented uses an approach that is based on patches. The intensity of one pixel is compared to the intensity of all of the other pixels in the image to generate a similarity map using this approach. In this particular investigation, the authors obtained edge information by using a Gaussian filter.

2.4. Literature Survey of Classification Techniques. Unsupervised learning-based neural networks were proposed by Goswami and Bhaiya [23] for the categorization of brain tumors. The first step of the brain tumor diagnosis technique is picture preprocessing, and the second phase is the extraction of tumors from MR images. Equalization of histograms, edge detection, noise filtering, and thresholding were all used in the picture preprocessing procedure. Independent component analysis was used to extract the brain feature (ICA). The self-organized map is used in the third phase to diagnose brain cancers (SOM). The brain scans were segmented using the K-means technique. Classification based on unsupervised learning was shown to be promising in the segmentation of brain tumor pictures in the aforementioned investigation.

Brain cancers may be detected using an automated approach [24]. A multi-stage tumor extraction procedure was used to identify the brain tumor automatically. The MR scans of the brain have the noise reduced. Features were then retrieved from noise-free brain scans, as seen in the figure. It was based on the extracted characteristic that classified the brain tumors. Brain tumors were classified using ensemble-based SVM. With the SVM-based classification approach, it was able to attain a 99 percent accuracy rate. The classification method uses multi-step segmentation techniques to identify the tumor in the brain MR images. For example, after removing the skull, FCM clustering methods were used to retrieve the afflicted region's brain tumors for inclusion in the algorithm.

The SOM clustering provided by Vaishnavee and Amshakala [25] was used to segment the brain pictures. Before the pictures are segmented, histogram equalization is used to extract the features. For the selection of features and to increase the accuracy of the classifiers, principal component analysis (PCA) was utilized. Proximal support vector machines (PSVM) classifiers were also developed, which were more successful than the SVM classifiers. According to our findings, the SVM classifier is a powerful classifier for extracting features from digital photos and other visual data.

3. Methodology

Using MRI images, this part presents machine learning and medically enabled multimodal for the segmentation and classification of brain tumors. MRI pictures contain noise. The geometric mean filter is utilized during picture preprocessing to facilitate the removal of noise. Fuzzy c-means algorithms are responsible for segmenting an image into smaller parts. The identification of a region of interest is facilitated by segmentation. The GLCM is utilized in order to carry out the process of dimension reduction. The GLCM algorithm is used to extract features from photographs. The photos are then categorized using various machine learning methods, including SVM, RBF, ANN, and AdaBoost. This model is shown in Figure 2.

Image preprocessing makes it feasible for illnesses seen in photographs to be classified in a more specific manner. Noise is the most prevalent type of picture artefact that can be observed in MRI scans; nevertheless, there is a broad range of image artefacts that can be noticed. Utilizing a number of distinct image filtering methods makes it possible to eliminate these artefacts from the image. In order to reduce the amount of noise present in the input photographs, a filter based on the geometric mean is applied to each of the images [26].

Clustering is a method that groups together patterns that are similar in order to discover the underlying connections that exist between individual pixels in an image. The process of putting things into groups or clusters based on the characteristics they have in common is what the term "clustering" refers to. When using the FCM methodology, the data items are sorted into groups according to the membership values of each category. Before splitting the final data, it is crucial to use the least squares strategy in

FIGURE 2: Machine learning and medically enabled multimodal for segmentation and classification of brain tumor using MRI images.

order to maximize the effectiveness of the object function [27].

Using a method known as feature extraction, which is part of the field of image processing, it is possible to exclude dimensions from a collection of feature subsets that are thought of as being unnecessary or unimportant. When texture properties need to be restored and links between pixels need to be maintained, the GLCM technique is called for. One way to go about doing this is by computing the co-occurrence values of the various grey levels. In order to construct the general linear model (GLM), conditional probability density functions known as $pIj|d$ and ş are utilized. After the GLM has been constructed, it is tested using distances d that range from one to five times the given direction (ş $= 0$, 45, 90, or 135). In order to accomplish this goal, the GLCM algorithm is used. According to the illustration, the probability that two pixels with the same grey level (I and/or j) as well as the inter-sample distance (d and ş) are spatially connected may be determined by applying the formula $p(i, j|d$, and ş). This formula also provides information about the distance between the samples (d). Contrast, correlation, entropy, and homogeneity are some of the most essential features of the general linear correspondence model (GLCM) [28].

This method, which is also known as AdaBoost, may be implemented to increase the accuracy of classification results produced by classifiers that are not all that great. This technique, known as AdaBoost, is used in the process of allocating initial weights to each observation. After a few iterations, incorrect categorizations will be given greater weight, while proper categorizations will be given less weight

as the results are iterated over. The weights that are given to each observation are determined by the classification to which it belongs. This was done so that the performance of the classifier might be improved. Because of this, the possibility of incorrect categorization is reduced. A large number of kids who are having difficulty academically are progressively fitted in an adjustable fashion using the process known as "boosting." In each consecutive model [29] in the series of models, there is a greater emphasis on information that had been ignored in earlier models.

In the realm of medicine, artificial neural networks (ANNs) are often used for the purpose of classifying medical images in order to arrive at a diagnosis. In many respects, including the manner in which it performs its duties, the ANN is comparable to the human brain. It is feasible to get the information required to make an informed guess about the category that an image belongs to by looking at a collection of photos that have already been categorized. This may be accomplished by looking at a collection of photographs that have been categorized. This may be accomplished by searching through a database of photos that have been arranged into a variety of different classifications. Every picture in this collection belongs to one of the categories that are listed above. ANNs are constructed by artificial neurons, each of which is programmed to perform in a manner that is analogous to the biological neurons found in the human brain. Connections allow for communication to take place between neurons that are located outside of the body. During the course of the learning process, weights may be given to neurons and edges, and those weights can be changed at any time according to the requirements of the task at hand. The standard structure of an artificial neural network consists of three layers: an input layer, a hidden layer, and an output layer that is responsible for signal generation. The majority of buildings will have this particular architectural style. Although the most typical topologies for artificial neural networks consist of input, hidden, and final layers, it is possible to configure the network in a variety of other ways. In theory, there may just be one hidden layer, there may be many hidden levels, or there may not be any hidden levels at all. It is within one's power to bring any one of these possibilities into fruition and make it a reality. In the event that it is required, the weights on a lower layer may be adjusted until the desired outcome is attained [30].

When it comes to the design of symbols, support vector machines provide a method that can discern the difference between the conceptions of measurement held by youngsters and those held by adults [31]. The following is a concise summary of the grouping issues faced by SVMs: (a) a nondirect translation of the information space to the higher measuring gimmick space, and (b) the construction of the distinguishing hyperplane that has the greatest distance from the purposes that are most closely related to the training set. Due to the direct distinct information, the SVM works toward locating, among all hyperplanes that minimize the preparation lapse, the particular case that distinguishes the preparation information with the greatest extreme separation from their nearest points of separation. This is
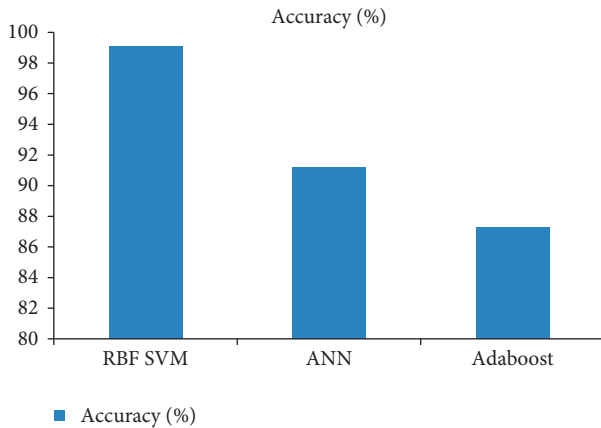
FIGURE 3: Accuracy comparison of classifiers for brain tumor detection.
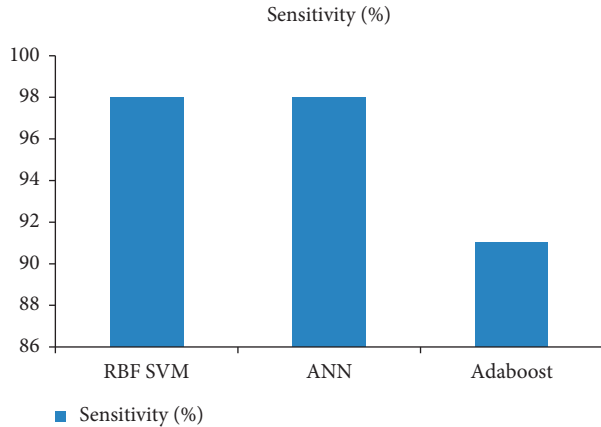
FIGURE 4: Sensitivity comparison of classifiers for brain tumor detection.
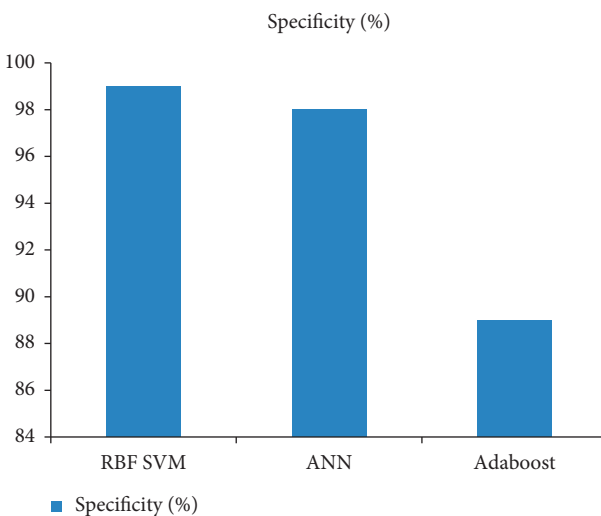
FIGURE 5: Specificity comparison of classifiers for brain tumor detection.

done in an effort to minimize the amount of time that is lost during the preparation process.

4. Result Analysis and Discussion

For the purpose of this investigation, one hundred pictures were chosen at random from Ref. [32]. Only 25 of the photos have tumors, while the other 75 are healthy and unaffected. In order to train the model, 80 photographs are used, and then 20 of those images are used for testing. The geometric mean filter is utilized during picture preprocessing to facilitate the removal of noise. Fuzzy c-means algorithms are responsible for segmenting an image into smaller parts. The identification of a region of interest is facilitated by segmentation. The GLCM is utilized in order to carry out the process of dimension reduction. The GLCM algorithm is used to extract features from photographs. The photos are then categorized using various machine learning methods, including SVM, RBF, ANN, and AdaBoost.

Within the scope of this investigation, a wide variety of algorithmic approaches are dissected and contrasted in terms of the degrees to which they exhibit accuracy, sensitivity, and specificity. Figures 3, 4, and 5 present the results of the classification performed by the algorithm. When it comes to the classification of brain tumors, the SVM RBF algorithm possesses an unrivaled level of accuracy and specificity that cannot be matched by any other method.

5. Conclusion

The abnormal and uncontrolled multiplication of brain cells may sometimes result in the development of a brain tumor. The majority of people who pass away from cancer do so because of brain tumors. As a direct result of this, it is becoming more difficult to pinpoint a therapy that is beneficial for a particular kind of brain tumor. Using a regular MRI scan, a three-dimensional picture of the human brain is possible to get. The major purpose of this organization is to investigate, recognize, diagnose, and categorize a wide range of neurological diseases. Radiation therapy is often used in the treatment of cancers, and MRI segmentation is frequently used in order to direct radiation therapy. As a result of this, we are in a position to determine whether or not a fragment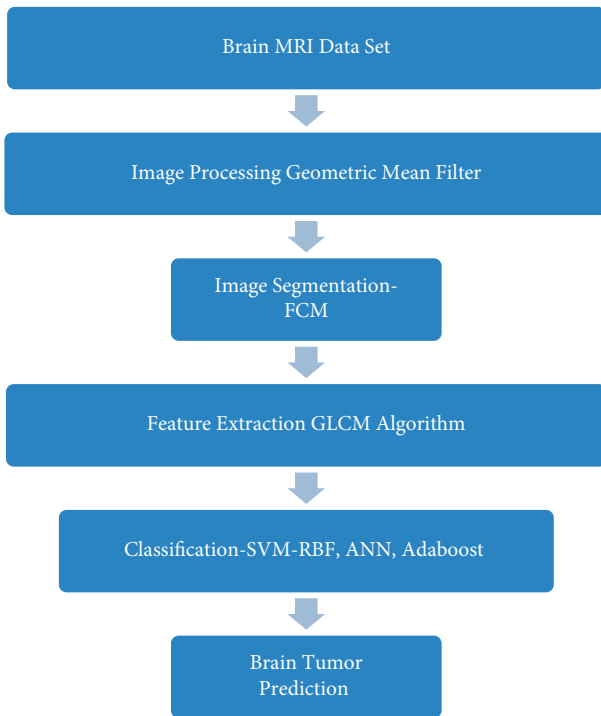 that was detected by an MRI is a tumor. Using MRI scans, this study offers a machine learning and medically enabled multimodal for segmentation and classification of brain tumors. MRI images contain noise. The geometric mean filter is used during picture preprocessing to facilitate the removal of noise. Fuzzy c-means algorithms are responsible for segmenting an image into smaller parts. The identification of an area of interest is facilitated by segmentation. The GLCM is used in order to carry out the process of dimension reduction. The GLCM algorithm is used to extract features from photographs. The photos are then categorized using several machine learning methods, including SVM, RBF, ANN, and AdaBoost. VM with RBF kernel is better for brain tumor detection using MRI images.

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] S. U. Aswathy, G. Glan Devadhas, and S. S. Kumar, "Brain tumor detection and segmentation using a wrapper based genetic algorithm for optimized feature set," *Cluster Computing*, vol. 22, no. S6, pp. 13369–13380, 2018.

[2] R. Anjali and S. Priya, "An efficient classifier for brain tumor classification," *IJCSMC*, vol. 6, no. 8, pp. 40–48, 2017.

[3] A. Raghuvanshi, U. K. Singh, and C. Joshi, "A review of various security and privacy innovations for IoT applications in healthcare," *Advanced Healthcare Systems*, pp. 43–58, 2022.

[4] V. D. P. Jasti, A. S. Zamani, K. Arumugam et al., "Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis," *Security and Communication Networks*, vol. 2022, Article ID 1918379, 7 pages, 2022.

[5] M. Angulakshmi and G. G. Lakshmi Priya, "Brain tumor segmentation from MRI using superpixels based spectral clustering," *Journal of King Saud University–Computer and Information Sciences*, vol. 32, no. 10, pp. 1182–1193, 2018.

[6] A. Sukumaran, D. G. Glan, and S. S. Kumar, "An improved tumor segmentation algorithm from T2 and FLAIR multimodality MRI brain images by support vector machine and genetic algorithm," *Cogent Engineering*, vol. 5, no. 1, Article ID 1470915, 2018.

[7] S. Chaudhury, A. N. Krishna, S. Gupta et al., "Effective image processing and segmentation-based machine learning techniques for diagnosis of breast cancer," *Computational and Mathematical Methods in Medicine*, vol. 2022, Article ID 6841334, 6 pages, 2022.

[8] S. González-Villà, A. Oliver, Y. Huo, X. Lladó, and B. A. Landman, "Brain structure segmentation in the presence of multiple sclerosis lesions," *NeuroImage: Clinic*, vol. 22, Article ID 101709, 2019.

[9] S. Agrawal and R. Sahu, "Wavelet based MRI image denoising using thresholding techniques," *International Journal of Science, Engineering and Technology Research (IJSETR)*, vol. 1, no. 3, pp. 28–33, 2011.

[10] S. S. Bedi and R. Khandelwal, "Arious image enhancement techniques- A critical review," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 3, pp. 1605–1609, 2013.

[11] S. J. Mathen and A. George, "Analysis of MRI enhancement techniques for contrast improvement and denoising," *International Journal of Current Engineering and Technology*, vol. 4, no. 6, pp. 3853–3861, 2014.

[12] A. D. Dina, S. A. Samy, and M. S. Ghoniemy Gamal, "Automated brain tumor detection and identification using image processing and probabilistic neural network techniques," *International Journal of Image Processing and Visual Communication*, vol. 1, no. 2, pp. 1–8, 2012.

[13] G. A. Leela and H. M. Veenakumari, "Morphological approach for the detection of brain tumor and cancer cells," *Journal of Electronics and Communication Engineering Research*, vol. 2, no. 1, pp. 7–12, 2014.

[14] J. Rohini Paul, C. Senthil Singh, and M. Manikandan, "Brain tumor MRI image segmentation and detection in image processing," *International Journal of Renewable Energy Technology*, vol. 3, no. 1, 2014.

[15] S. Mohammed, Al-T. Hamoud, and G. Sulong, "Tumor brain detection through MR images: a review of literature," *Journal of Theoretical and Applied Information Technology*, vol. 62, no. 2, 2014.

[16] S. Siva Sankari, M. Sindhu, R. Sangeetha, and A. ShenbagaRajan, "Feature extraction of brain tumor using MRI," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 3, 2014.

[17] K. K. Reddy, B. Solmaz, P. Yan, N. G. Avgeropoulos, D. J. Rippe, and M. Shah, "Confidence Guided Enhancing Brain Tumor Segmentation in Multi-Parametric MRI," in *Proceedings of the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 366–369, IEEE, Barcelona Spain, May 2012.

[18] M. Soni, A. Khare, and S. Jain, "A survey of digital image processing and its problem," *International Journal of Scientific and Research Publications*, vol. 4, no. 2, 2014.

[19] C. Hemalatha, S. Muruganand, and R. Maheswaran, "Preprocessing methods to remove impulse noise in avian pox affected Hen Image using Image Processing," *International Journal of Computer Application*, vol. 98, no. 20, pp. 18–21, 2014.

[20] P. Mohanaiah, P. Sathyanarayana, and L. GuruKumar, "Image textural feature extraction using GLCM approach," *International Journal of Scientific and Research Publications*, vol. 3, no. 5, pp. 1–5, 2013.

[21] D. Ping Tian, "A review on image feature extraction and representation techniques," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 8, no. 43, pp. 85–396, 2013.

[22] K. Jafari-Khouzani, "MRI up sampling using feature-based nonlocal means approach," *IEEE Transactions on Medical Imaging*, vol. 33, no. 10, pp. 1969–1985.

[23] S. Goswami and L. K. P. Bhaiya, "Brain Tumor Detection Using Unsupervised Learning Based Neural Network," in *Proceedings of the International Conference on Communication Systems and Network Technologies (CSNT)*, pp. 573–577, IEEE, Gwalior India, April 2013.

[24] N. B. Bahadure, A. K. Ray, and H. P. Thethi, "Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM," *International Journal of Biomedical Imaging*, vol. 2017, Article ID 9749108, 12 pages, 2017.

[25] K. B. Vaishnavee and K. Amshakala, "An automated MRI brain image segmentation and tumor detection using SOM-clustering and Proximal Support Vector Machine classifier," in *Proceedings of the 2015 IEEE International Conference on Engineering and Technology (ICETECH)*, pp. 1–6, Coimbatore India, March 2015.

[26] M. Kalhor, A. Kajouei, F. Hamidi, and M. M. Asem, "Assessment of histogram-based medical image contrast enhancement techniques; an implementation," in *Prcoeeding of the 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0997–1003, Las Vegas NV USA, January 2019.

[27] V. V. Vela-Rincón, D. Mújica-Vargas, M. Mejía Lavalle, and A. Magadán Salazar, "Spatial $\alpha\alpha$-trimmed Fuzzy C-means algorithm to image segmentation," in *Pattern Recognition. MCPR 2020*, K. Figueroa Mora, J. Anzurez Marín, J. Cerda, J. Carrasco-Ochoa, J. Martínez-Trinidad, and J. Olvera-López, Eds., vol. 12088, Cham, Springer, 2020 Lecture Notes in Computer Science.

[28] M. Benco, P. Kamencay, M. Radilova, R. Hudec, and M. Sinko, "The comparison of color texture features extraction based on 1D GLCM with deep learning methods," in *Proceedings the International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 285–289, IEEE, Niteroi Brazil, July 2020.

[29] T. R. Mahesh, V. Dhilip Kumar, V. Vinoth Kumar et al., "AdaBoost Ensemble methods using K-fold cross validation for survivability with the early detection of heart disease," *Computational Intelligence andNeuroscience*, vol. 2022, Article ID 9005278, 11 pages, 2022.

[30] M. A. Khan and F. Algarni, "A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO-ANFIS," *IEEE Access*, vol. 8, pp. 122259–122269, 2020.

[31] E. F. Ohata, G. M. Bezerra, J. V. S. Chagas et al., "Automatic detection of COVID-19 infection using chest X-ray images through transfer learning," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 1, pp. 239–248, 2021.

[32] Sartaj, "Brain tumor classification (MRI)," 2020, https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri.

Retraction

Retracted: The Influence of Martial Arts on Spine CT Image Morphological Structure Based on Optimized Ant Colony 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Yuhuan, H. Xueting, and W. Weiyue, "The Influence of Martial Arts on Spine CT Image Morphological Structure Based on Optimized Ant Colony Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2725819, 10 pages, 2022.

Research Article

The Influence of Martial Arts on Spine CT Image Morphological Structure Based on Optimized Ant Colony Algorithm

Hou Yuhuan [ID],[1,2] Hou Xueting [ID],[1,3] and Wang Weiyue [ID][3,4,5]

[1]*Basic Medicine Department, Hubei College of Chinese Medicine, Jinzhou, Hubei 434000, China*
[2]*Dankook University College of Physical Education, Yongin, Gyeonggi Province 330712, Republic of Korea*
[3]*College of Life Sciences, Guangxi Normal University, Guilin, Guangxi 541000, China*
[4]*Shenyang Sports Universitye, Shenyang, Liaoning 110102, China*
[5]*Wuhan Sports University, Wuhan, Hubei 430079, China*

Correspondence should be addressed to Wang Weiyue; 161847331@masu.edu.cn

As the second lifeline of human body, the abnormal development of its morphological structure has a great impact on people's physical and mental health. Due to bad living habits and learning pressure, the morphological and structural development of spine in adolescents and children is abnormal to a certain extent, which can be improved by sports intervention. Therefore, this article puts forward the research on the influence of Wushu on the spinal morphological structure based on the optimization algorithm and integrates the optimization ant colony algorithm on the basis of the traditional spinal CT image segmentation method. The experimental results show that the improved CT image segmentation method based on the optimized ant colony algorithm can solve the sensitive problem of the number of clusters, improve the segmentation efficiency and quality, and provide more accurate data for the subsequent comparative experiments. At the same time, the comparative test results show that Wuqinxi Wushu can better improve the abnormal dry tilt angle, abnormal kyphosis angle, and body balance angle of teenagers, improve the activity of teenagers' spine, and help teenagers enhance the overall health of spine.

1. Introduction

Spine is the second lifeline of human body, and its abnormal development has a negative impact on human life. The key period of spinal development is in human adolescence, which is also the golden period of physical development of adolescents and children. With the growth of age, the height and weight of teenagers will continue to increase. If there are nutritional imbalance and lack of exercise at this time, coupled with the low content of inorganic salts in the bones of teenagers and children, under the condition of relatively low bone hardness, teenagers' spine is easy to be deformed under the influence of bad living habits or gravity compression of the external environment, resulting in abnormal spine development [1]. With the development of information technology, while the academic pressure of teenagers has been increasing in recent years, the homework and curriculum tasks completed through the information network are also increasing. Many teenagers' spine has been deformed due to long-time desk writing, poor sitting posture, long-term bow, and heavy schoolbag pressure, even some teenagers and children have spinal lesions, and the prevalence of spine is increasing every year [2]. Therefore, scientific and effective intervention research on spinal development of adolescents and children has become a hot spot in the society.

In the past, relevant experts focused more on the abnormal spinal morphology of adults and the elderly and paid less attention to the abnormal spinal morphology and structure of adolescents and children [3]. However, with the increasing incidence rate of spine and younger age of onset, the spine research is deepening and attention is being paid to the development of spinal column in young children. According to relevant research experience,

sports is the simplest means to correct spine shape. Some scholars use Yoga treatment to treat female college students with scoliosis [4]. The quality results show that Yoga treatment has significantly improved scoliosis and shoulder balance, and the mobility of spine in all aspects has also been improved; it shows that Yoga treatment has an effect on spinal correction [5]. Some scholars conducted comprehensive intervention on middle school students with abnormal spinal development for two months, and through the comparison of spinal morphology measurement results before and after the experiment, it is concluded that the degree of spinal abnormality has been greatly reduced, and more than half of the students' spinal posture can return to the normal state, which has a significant improvement effect on spinal abnormality [6]. Other scholars used Taijiquan to treat the middle-aged and elderly patients with lumbar pain symptoms for six months [7]. The chronic low back pain symptoms of the experimental subjects were relieved, and the lumbar spine function was improved [8]. Based on this, some scholars have proposed to effectively intervene the spine morphological development of adolescents and children through martial arts science, so as to help teenagers and children maintain and improve the healthy state of spinal development [9].

However, at present, there is relatively little research on the impact of martial arts on spinal morphological development [10]. Therefore, this paper proposes the impact of martial arts on spinal morphological structure based on the optimized ant colony algorithm [11]. Through the improved CT image segmentation method of the optimized ant colony algorithm, this paper analyzes all aspects of the experimental object's spinal CT images before and after the experiment [12]. This paper is mainly divided into three parts [13]. The first part expounds the influencing factors of spinal morphological structure development and the harm of abnormal development; the second part is the construction of the spine development evaluation model based on optimized ant colony; the third part is the experimental results and analysis of the influence of spine shape and structure of Wushu team based on the optimized ant colony algorithm [14].

The innovation of the research lies in the integration of the optimized ant colony algorithm on the basis of traditional spine CT image segmentation methods. This algorithm can solve the sensitive problem of clustering number, improve the efficiency and quality of segmentation, and provide more accurate data for subsequent comparative experiments. Compared with the traditional fuzzy c-means classification method, the improved CT image segmentation method based on the optimized ant colony algorithm solves the problem of being sensitive to the number of clusters and improves the segmentation efficiency while ensuring the quality of image segmentation. Based on the optimization algorithm, the influence of Wushu on the spine shape and structure was analyzed. It can improve the abnormal development of teenagers' spine and help teenagers improve their overall health.

2. Factors Affecting the Development of Spinal Morphology and Structure and the Harm of Abnormal Development

The human spine is located in the middle of the back of the human body and is the backbone of the human body [15]. The spine has a central axial bone structure, which plays a role of lever and supporting weight-bearing in the human body [16]. The physiological bending of the spine can improve the elasticity of the spine and buffer and disperse the gravity and impact on the intracranial central nerve [17]. The flattening of the intervertebral disc in the spine can help the spine to achieve movement and complete complex movements and expand the range of motion of the spine [18]. As can be seen, spine has a very important impact on human activities [19]. Abnormal spinal morphological development has a serious adverse impact on the life and learning of teenagers and children [20]. Figure 1 shows a comprehensive view of the spine.

There are three main influencing factors in the development of spine, namely, genetic factors, lifestyle and habits, and exercise factors [21]. Common spinal diseases such as mandatory spondylitis and adolescent idiopathic scoliosis are affected by genetic genes. Although genetic gene is a very important factor in the development of spine in adolescents and children, it is not a decisive factor. The acquired development of spine is also affected by the living habits and styles of adolescents and children, including positive and negative effects [22]. According to the relevant research results, the quality and manner of teenagers' backpacks will affect the development of their spine morphology, that is, the heavier the backpack is, the more obvious the morphological changes of teenagers' spine will be, which improves the possibility of spine damage [23]. At the same time, teenagers' bad habits such as unilateral backpacking and long-term incorrect sitting posture will increase the burden on the spine, affect the activity and development health of different parts of the spine, and improve the probability of spinal disease. In addition, according to relevant studies, reasonable and appropriate exercise will have a positive impact on the development and morphological structure of teenagers' spine, while different exercise modes and mild exercise have different effects on the spine. Therefore, in the process of helping teenagers and children maintain and improve spinal development and morphological structure, we should find appropriate sport methods.

Abnormal development of spinal morphology and structure will not only bring physical harm to teenagers and children but also have a negative impact on their psychology. According to relevant statistical data, the number of Chinese teenagers suffering from scoliosis of different degrees accounts for 20% of the total number of teenagers and children. More than 80% of teenagers have poor physical posture, of which the most common and serious problem is poor sitting posture, neck probing, and chest hunchback when walking [24]. Such a situation is not only conducive to physical and mental health in the process of teenagers' growth but also has a serious impact on their behavior in
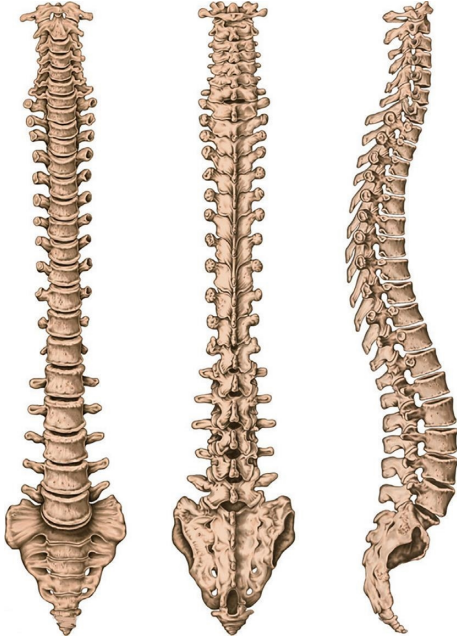
Figure 1: Comprehensive view of spine.

future study and life. The abnormal development of spinal morphology and structure is very harmful to the health of adolescents. For example, when the most common scoliosis is in a serious state, it will damage lung function, produce waist and back pain, and affect the biological stress balance of spine [25]. If we cannot take effective measures to intervene and treat in time, it will not only aggravate the condition but also increase the difficulty of treatment.

3. Construction of Spine Development Evaluation Model Based on Optimized Ant Colony Algorithm

Due to the gradual aging of the lumbar spine in the middle-aged and elderly, scoliosis can be found in the frontal view. From the lateral position, the disappearance of physiological protrusion may occur, and the disappearance of cervical physiological protrusion may occur in the cervical spine, or spinal deformity is caused by degenerative diseases such as cervical spondylosis. Adolescent spinal deformities, such as scoliosis, can be divided into idiopathic scoliosis, congenital scoliosis, and scoliosis secondary to the muscular nervous system. With the development of science and technology and computer technology, the medical means of analyzing the abnormal development and disease causes of human spine are also constantly developing and updating. Among them, biological realistic model is a method that can present the true state of spine, which integrates computer graphics, digital image processing, medical image knowledge, and other disciplines. It mainly carries out three-dimensional reconstruction of the biological realistic model based on spinal CT images. It is one of the methods to judge the

morphological structure of spinal development. Image segmentation is an important part of image 3D reconstruction. The quality of image segmentation is directly related to the accuracy of image analysis, understanding, and 3D reconstruction. Medical image segmentation is to segment the required regions in the image according to the differences between regions and the similarity within regions. At the same time, it should ensure that there is no overlap between the regions of the segmented image. Let all the region sets of the image be represented as R, and nonempty region segmented by it be represented as $R_1, R_2, \ldots, R_n$. There are four conditions to be met. The first one is shown in the following formula:

$$R = \cup_{i=1}^{n} R_i. \tag{1}$$

Second, the two random subregions need to meet the conditions, as shown in the following formula:

$$R_i \cap R_j = \varnothing, \tag{2}$$

in which $i \neq j$. Third, all subregions are nonempty, that is, $R_i = \varnothing, i = 1, 2, \ldots, n$; fourth, all subregions are connected regions.

3.1. CT Image Segmentation Method Based on Fuzzy C-Means Clustering. Fuzzy C-means clustering method is a clustering algorithm that uses the membership function to determine the clustering degree of pixels. Its core idea is to divide the objects with the greatest similarity into the same cluster, and the similarity between different clusters is the smallest. Let the vector be expressed as X_i and $i = 1, 2, \ldots, n$, and its number is n. Divide it into fuzzy subsets with the number of c and $2 \leq c \leq n$. The membership degree of sample points is expressed as μ_{ik}, and the classification results can be expressed by fuzzy membership degree matrix, as shown in the following formula:

$$U = \{\mu_{ik}\}, \tag{3}$$

in which $\mu_{ik} \in (0, 1)$. The membership degree is normalized to obtain the following formula:

$$\sum_{i=1}^{c} \mu_{ik} = 1, \tag{4}$$

in which $k \in [1, n]$.

The value function expression of fuzzy c-means clustering is shown in the following formula:

$$J_m(U, V: X) = \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik})^m \|X_k - X_i\|_A^2, \tag{5}$$

where the strength matrix is expressed as U, the positive definite matrix is expressed as A, the central point set of clustering fuzzy subset is expressed as $V = (v_1, v_2, \ldots, v_c)$, and the weighted index is expressed as $m \in [1, \infty)$. In order to solve the optimization of formula (5), formula (6) is obtained by constructing Lagrange functional, as follows:

$$\overline{J}(U, V: X) = J(U, V: X) + \sum_{j=1}^{n} \lambda_i \sum_{i=1}^{c} (\mu_{ik} - 1) = \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik})^m \|X_k - X_i\|_A^2 + \sum_{j=1}^{n} \lambda_i \left(\sum_{i=1}^{c} \mu_{ik} - 1 \right), \tag{6}$$

where the Lagrange multiplier is expressed as λ_i, $i = 1, 2, \ldots, n$. Calculate the partial derivative of all functional variables used in formula (6) to achieve the minimum necessary condition for realizing the optimal solution, as shown in the following formulas:

$$c_i = \frac{\sum_{j=1}^{n} \mu_{ij}^m x_j}{\sum_{j=1}^{n} \mu_{ij}^m}, \tag{7}$$

$$\mu_{ij} = \frac{1}{\sum_{K=1}^{n} \left(d_{ij}/d_{kj} \right)^{2/(m-1)}}, \tag{8}$$

where the Euclidean distance between the j point and the i cluster center is expressed as d_{ij}.

3.2. Improved CT Image Segmentation Method Based on Optimized Ant Colony Algorithm.

The image segmentation by the fuzzy c-means clustering method has good quality and retains the complete information of the required region, but the segmentation time is relatively long. In order to reduce the number of iterations and improve the efficiency of segmentation, this paper introduces the optimal ant colony clustering algorithm.

Ant individuals are very simple social insects, but when they carry out group activities, they can complete extremely complex behaviors through the cooperation between individuals, showing the intelligence in their social tasks. An ant colony can find the shortest path between the nest and food when looking for food and can make corresponding adjustments in a short time when the external environment changes. Ant colony can complete these behaviors, which mainly rely on information hormone, which can transmit the signal to be expressed by individual ants to the latter, so as to affect the behavior of the latter. In addition, if there are obstacles on the way of ant colony foraging, the individual ant choosing the shortest path will release information hormone and affect the behavior of the following ants. With the higher content of information hormone, as shown in Figure 2.

The optimized ant colony image segmentation algorithm mainly includes the characteristics of systematicity, distributed computing, self-organization, and positive feedback. As shown in Figure 3, it is an improved CT image segmentation flow chart integrating the optimized ant colony algorithm. The ant colony clustering algorithm is used to extract the initial cluster center of the image to be segmented, and then the fuzzy c-means algorithm is used to segment the image based on the initial cluster center.

Let the image be expressed as X, the pixel X_i in the image is regarded as an ant individual, and each ant has three directional vectors of gradient, gray, and domain features. The distance between the pixel and the cluster center is

expressed as d_{ij}, which is calculated by Euclidean distance, as shown in the following formula:

$$d_{ij} = \sqrt{\sum_{k=1}^{m} p_k \left(X_{ik} - C_{jk} \right)^2}. \tag{9}$$

The basic characteristic dimensions of ants are m and $m = 3$, and the weighted silver of each dimension is p_k.

The amount of information contained in the area is expressed as ph_{ij}, the influence range of ant individuals on the amount of information in the surrounding local area is expressed as r, and the influence degree is shown in the following formula:

$$ph_{ij} = \begin{cases} 1, d_{ij} \leq r; \\ 0, otherwise. \end{cases} \tag{10}$$

When the ant individual makes path selection, let the attraction factor from the current pixel to the cluster center at the t iteration be expressed as $p_{ij}^k(t)$, and its expression is shown in the following formula:

$$p_{ij}^k t = \begin{cases} \dfrac{\left[p_{ij}^k t \right]^\alpha \left[\eta_{ij} t \right]^\beta}{\sum_{s=allowed_k} \left[p_{ij}^k t \right]^\alpha \left[\eta_{ij} t \right]^\beta}, if \ j \in allowed_k; \\ \\ 0, \ or \ other \end{cases} \tag{11}$$

The residual information factor is α, the heuristic information factor is β, and the set of ant feasible paths is $allowed_k$, as shown in the following formula:

$$allowed_k \in \left\{ X_s | d_{sj} \leq r, s = 1, 2, \ldots, N \right\}. \tag{12}$$

The heuristic function is expressed as $\eta_{ij}(t)$, and its formula is as follows:

$$\eta_{ij} t = \frac{r}{d_{ij}} = \frac{r}{\sqrt{\sum_{k=1}^{m} p_k \left(X_{ik} - C_{jk} \right)^2}}. \tag{13}$$

As can be seen from formula (13), the heuristic function increases with the increase of cluster radius, that is, the probability of selecting the corresponding cluster center increases. In addition, the heuristic function decreases with the increase of the distance between the pixel and the cluster center, that is, the probability of ant individuals selecting the corresponding cluster for merging decreases.

The main judgment basis for whether a pixel belongs to a cluster center is shown in the following formula:

$$p_{ij}^k(t) > p_0. \tag{14}$$

If formula (14) is not met, the pixels are not merged into the cluster.

After one iteration, all ants need to update the cluster center and intraclass dissimilarity. As shown in formulas
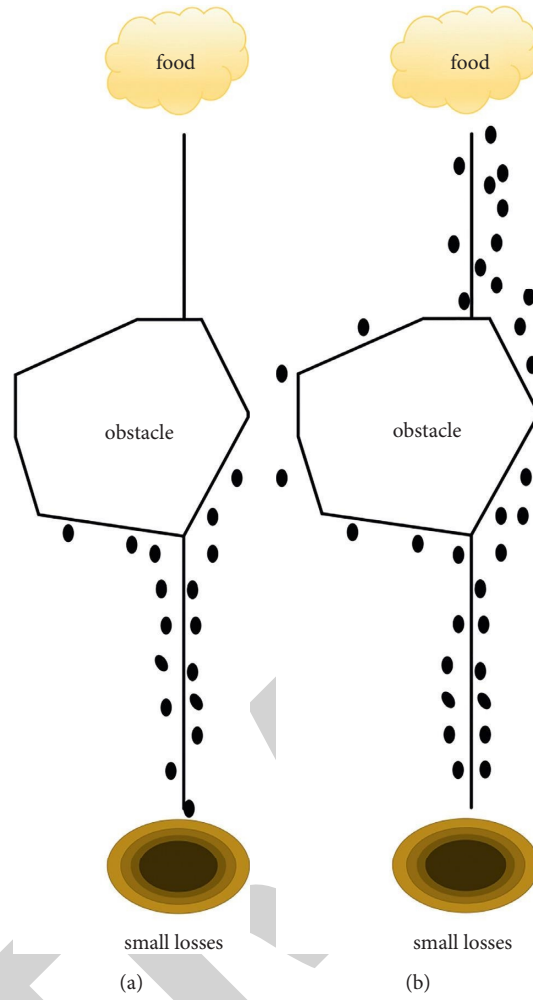
Figure 2: Schematic diagram of ant colony foraging behavior in natural biosphere.

(15) and (16), it is a new calculation method of cluster center and dissimilarity.

$$\widetilde{C}_j = \frac{1}{j} \sum_{k=0}^{j} (x_k), \tag{15}$$

where $k \in (1, 2, \ldots, j)$; the new cluster center is expressed as $\widetilde{C}_j$.

$$\widetilde{\varepsilon} = \sum_{j=1}^{k} \sum_{k=1}^{j} \sqrt{\sum_{i=1}^{m} \left(X_{ki} - C_{ji} \right)^2}. \tag{16}$$

The new dissimilarity is expressed as $\widetilde{\varepsilon}$.

4. Experimental Results of the Influence of Martial Arts on Spine Shape and Structure Based on Optimized Ant Colony Algorithm

In order to better test the improved CT image segmentation method based on the optimized ant colony algorithm, the segmentation time and iteration times are compared with the traditional fuzzy c-means classification method. First, the fuzzy rules are determined. Enter the corresponding parameters and activate some fuzzy rules. Only the rules whose membership degree is not 0 can be activated. The conclusion is obtained by comparing the membership degree, but the conclusion is not certain at this time. Finally, the conclusion is used to determine the actual output as shown in Figure 4.

It can be seen from the results in the figure that the time and iteration times of the two image segmentation methods increase with the increase of the number of clusters. Compared with the traditional fuzzy c-means classification method, the improved CT image segmentation method based on optimized ant colony algorithm is less sensitive to the number of clusters, and the increase of time and algorithm iteration is much less than the traditional fuzzy c-means classification algorithm. This shows that the improved CT image segmentation method based on optimized ant colony algorithm effectively overcomes the sensitivity of traditional methods to the number of clusters, improves the quality and efficiency of segmentation, and can provide more accurate information
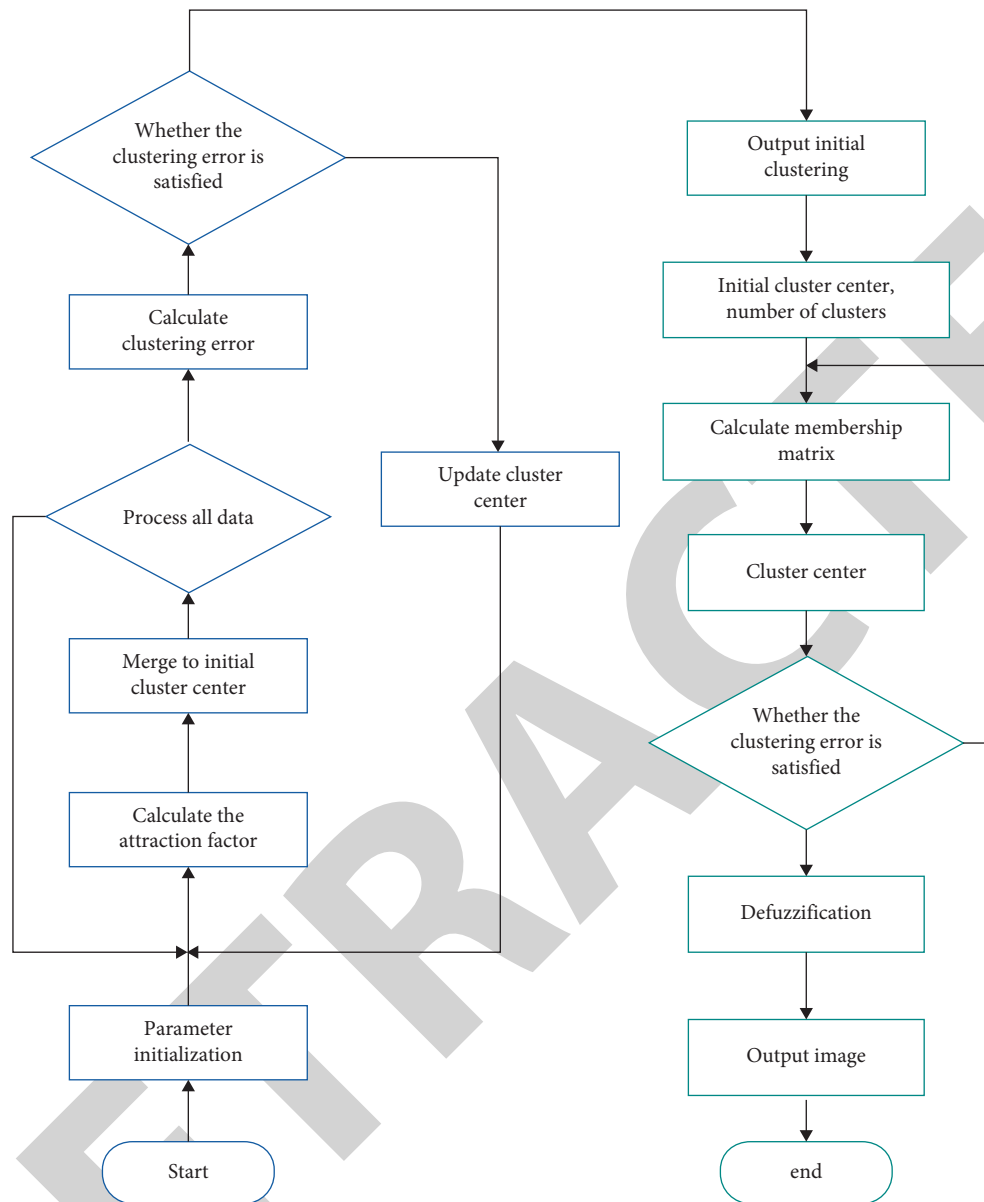
FIGURE 3: Improved CT image segmentation flowchart with optimized ant colony algorithm.

for the study of the impact of Wushu on the morphological structure of spine.

Huatuo Wuqinxi is an integral part of Chinese martial arts qigong, and it is also one of the national intangible heritages in Chinese traditional sports culture. The long-term development of Huatuo Wuqinxi is mainly because its movements are simple and easy to learn, suitable for all ages, and have a good mass foundation. Compared with other martial arts content, it can be more accepted and recognized by people. In this paper, 45 students with abnormal spinal morphology and structure in a primary school were selected as the experimental objects and randomly assigned to the experimental group and the control group. In the experimental group, eight subjects were patients with abnormal trunk inclination, fifteen subjects were patients with abnormal kyphosis angle, and

two of them were patients with abnormal two conditions at the same time. In the control group, eight subjects had abnormal trunk inclination, fourteen subjects had abnormal kyphosis angle, and one of the subjects had abnormal two conditions at the same time.

Figure 5 shows the comparison of trunk tilt angle measurement results of experimental subjects before and after the experiment between experimental group and control group. Before the experiment, there was no significant difference in trunk tilt angle between the experimental group and the control group, while after the experiment, there was no significant change in trunk tilt angle between the two groups, $P > 0.05$, that is, there was no significant difference. The comparison $P < 0.01$ between the experimental group before and after the intervention shows that the trunk tilt angle of the experimental group is better than

Figure 4: The traditional fuzzy c-means classification method is based on the improved CT image segmentation method of optimized ant colony algorithm.
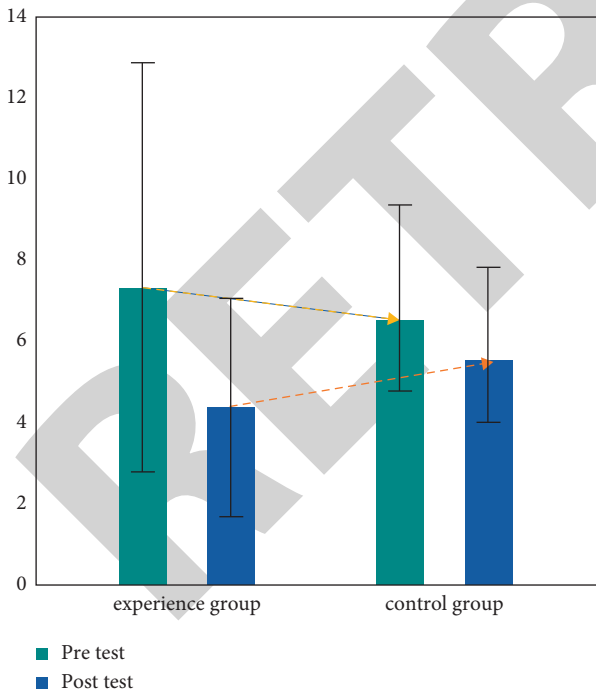
Figure 5: Comparison of trunk tilt angle measurement results between experimental group and control group before and after experiment.

that before the intervention of Wuqinxi, and there is a significant difference.

As shown in Figure 6, the results of kyphosis angle of subjects in the experimental group and the control group
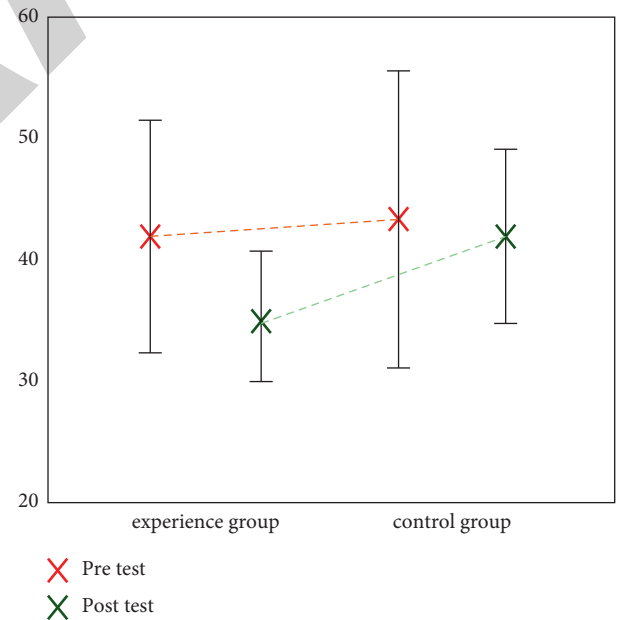
Figure 6: Comparison of kyphosis angle between experimental group and control group before and after experiment.

before and after the experiment are compared. It can be seen from the figure that before the intervention of Wuqinxi, there was no obvious difference in the kyphosis angle between the experimental group and the control group, that is, $P > 0.05$, while after the intervention of Wuqinxi, the kyphosis angle of the experimental group was significantly improved, forming a very obvious difference with the
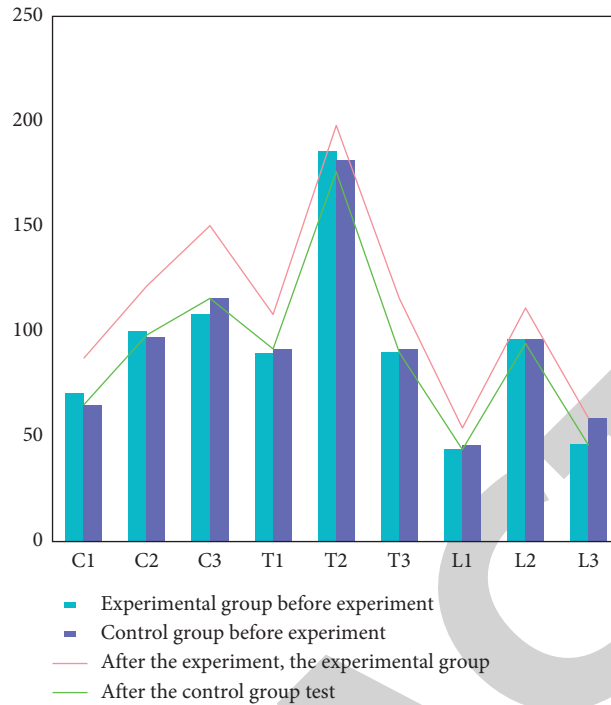
Figure 7: Comparison of measurement results of spinal mobility between experimental group and control group.
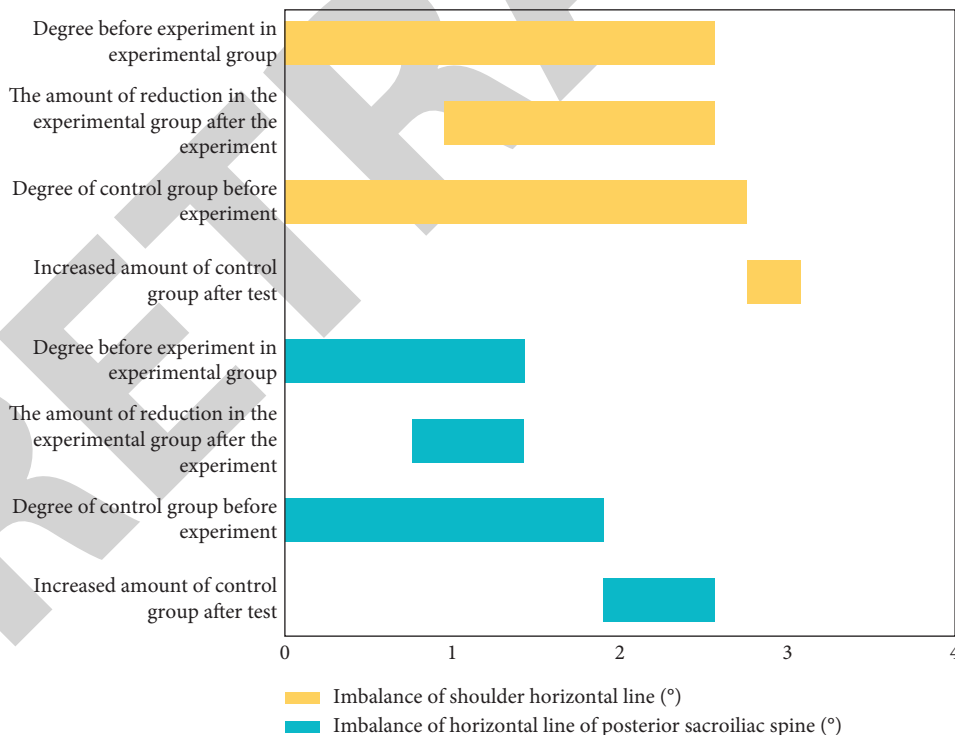
Figure 8: Comparison of body balance results between experimental group and control group.

control group. The kyphosis angle of the experimental group was significantly improved than that before the intervention of Wuqinxi, and there was significant difference between the two.

As shown in Figure 7, the comparison of spinal activity measurement results of subjects in the experimental group

and the control group is shown. It can be seen from the figure that there is no significant difference in spinal activity between the experimental group and the control group before the intervention of Wuqinxi. After the intervention of Wuqinxi, the cervical spine activity and thoracic spine activity of the experimental group were better than those of the

control group. There was no significant difference in lumbar spine activity between the experimental group and the control group. In addition, compared with before the intervention of Wuqinxi, the subjects in the experimental group improved their mobility in three directions: coronal, sagittal, and horizontal. The angle of motion of thoracic vertebrae was significantly improved in coronal and horizontal directions, and there were significant differences. The improvement of lumbar mobility in three directions is different, among which the improvement of lumbar mobility in sagittal plane is the most obvious and effective, with very significant difference. Second, there is a certain improvement in the activity of coronal plane, with significant difference. Finally, the activity of lumbar spine in the horizontal plane has a certain improvement trend, but there is no obvious difference.

Figure 8 shows the comparison of body balance results between the experimental group and the control group. It can be seen from the figure that before the intervention experiment of Wuqinxi, there was no significant difference in body balance between the two groups. After the intervention of Wuqinxi, there was a very significant difference between the experimental group and the control group. The measurement results of the experimental group were better than those of the control group. At the same time, the shoulder horizontal line of the experimental group was significantly improved compared with that before the intervention of Wuqinxi and showed a very obvious difference, namely, $P < 0.01$. The horizontal line of the posterior sacroiliac spine of the experimental group was also improved compared with the previous, with obvious differences.

In conclusion, the intervention of Huatuo Wuqinxi martial arts can improve the abnormal trunk tilt angle, kyphosis angle, and body balance angle of teenagers, and the effect is good. At the same time, Huatuo Wuqinxi can also improve the activity of teenagers' cervical spine and thoracic spine, showing obvious differences. There is also a certain improvement effect on the lumbar mobility. The improvement of the lumbar mobility in three aspects is different. The lumbar mobility is improved in the coronal plane and sagittal plane, showing a significant difference, but the improvement effect on the lumbar mobility in the horizontal plane is not obvious. Therefore, Huatuo Wuqinxi can improve the health status of teenagers' spine and improve the abnormal development of spine morphology and structure.

5. Conclusion

Adolescence and childhood are not only the golden periods for the development of human spine morphology and structure but also the key period for the physical development of adolescents and children. Therefore, adolescents' and children's spines are prone to abnormal development due to internal and external environment. At the same time, the bad living habits and learning pressure of teenagers have caused a certain pressure on the developing spine, which promotes its abnormal development. Most teenagers have slight spinal dysplasia such as chest hump and back compression, which can be improved through exercise

intervention. In Chinese traditional sports culture, Wuqinxi is one of the national intangible cultural heritages, which has always had good results in health preservation. Therefore, this paper puts forward the research on the influence of martial arts on the spinal morphological structure based on the optimized ant colony algorithm. By optimizing the ant colony algorithm, we can improve the segmentation method of spinal CT image, improve the quality and efficiency of image segmentation, and provide more effective information and data for the later analysis of spinal morphological structure. Experiments show that compared w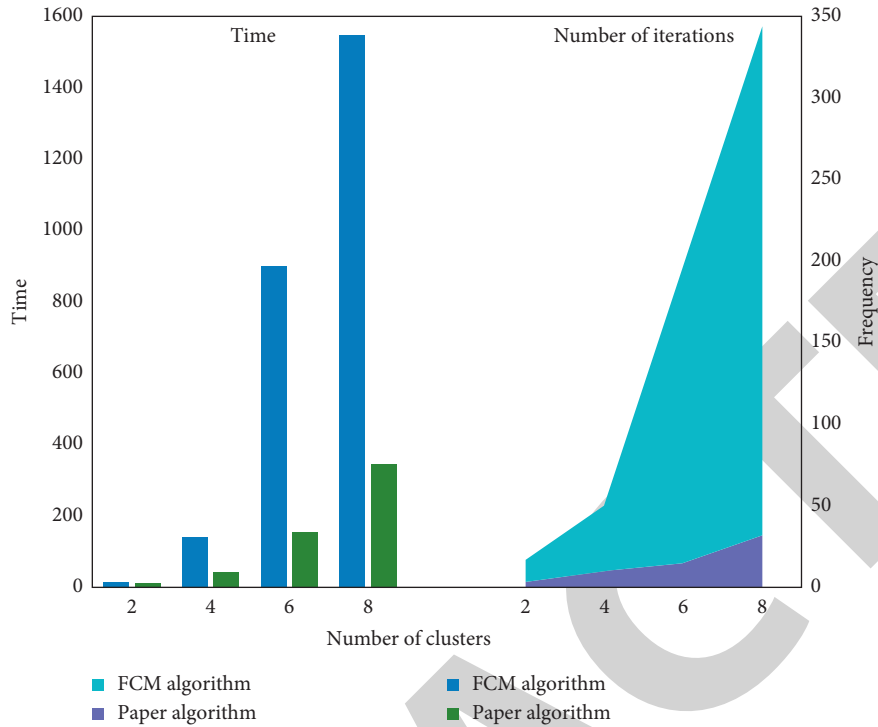ith the traditional fuzzy c-means classification method, the improved CT image segmentation method based on optimized ant colony algorithm solves the problem of being sensitive to the number of clusters and improves the segmentation efficiency while ensuring the image segmentation quality. The comparison of the experimental results between the experimental group and the control group shows that the intervention of Wuqinxi Wushu can improve the abnormal trunk tilt angle and kyphosis angle of teenagers and achieve better results. In terms of adolescent spinal mobility, Wuqinxi has a good effect on the mobility of cervical spine and thoracic spine in three directions and has a significant effect on the mobility of lumbar spine in coronal plane and sagittal plane, but it has no obvious improvement on the mobility of lumbar spine in horizontal plane. Therefore, Huatuo Wuqinxi Wushu can improve the abnormal development of teenagers' spine and help teenagers improve their overall health. However, the study still has some limitations. The number of subjects is too small, so whether it can improve the health of adolescent spine is still controversial. Whether it is representative to improve the abnormal development of spinal morphology and structure remains to be discussed.

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] F. Meng and B. Cui, "Analysis on the current situation and countermeasures of college students' extracurricular sports activities -- taking huaibei normal university as an example," *Bulletin of Sport Science & Technology*, vol. 27, no. 07, pp. 113–115, 2019.

[2] Q. Lou and A. Wang, "Effect of trunk strength training on correcting spinal curvature of table tennis players," *Contemporary sports technology*, vol. 39, no. 14, p. 33, 2019.

[3] Z. Jiang, H. Xu, and Y. Wan, "Research Progress on screening of abnormal spinal curvature in children and adolescents," *Chinese Journal of School Health*, vol. 42, no. 02, pp. 312–315, 2021, + 320.

[4] K. Zhou, Y. Chen, and Y. Cao, "Preliminary study on the related status of cervical sub-health," *Contemporary medicine*, vol. 2019, pp. 185–188, 2019.

Research Article

Influence of Autologous Bone Marrow Stem Cell Therapy on the Levels of Inflammatory Factors and Conexin43 of Patients with Moyamoya Disease

Liming Zhao [ID],[1] **Tianxiao Li** [ID],[1] **Bingqian Xue** [ID],[2] **Hao Liang** [ID],[2] **Shao Zhang** [ID],[2] **Ruiyu Wu** [ID],[2] **Gaochao Guo** [ID],[1] **Tao Gao** [ID],[1] **Yang Liu** [ID],[1] **Yuxue Sun** [ID],[1] and **Chaoyue Li** [ID][1]

[1]*Department of Neurosurgery, Zhengzhou University People's Hospital, Henan Provincial People's Hospital, Zhengzhou 450003, China*
[2]*Department of Neurosurgery, Henan University People's Hospital, Henan Provincial People's Hospital, Zhengzhou 450003, China*

Correspondence should be addressed to Chaoyue Li; lichaoyue@mjcedu.cn

Moyamoya disease is a medical condition that shows the typical characteristics like continuous and chronic thickening of the walls and the contraction of the internal carotid artery; as a result, the internal diameter of the artery gets narrowed. There are six phases of the disease ranging from I to VI (moyamoya vessels completely disappear, followed by the complete blockage of the arteries). Surgery is a commonly recommended treatment for the moyamoya disease. Our research study identifies the effect of autologous bone marrow stem cell therapy (ABMSCT) on the levels of inflammatory factors and Conexin43 (Cx43) protein in patients suffering from moyamoya. In our study, we have selected 52 moyamoya patients admitted to our hospital from 30 July 2019 to 10 February 2020. The control group (CG) was treated with superficial temporal artery to a middle cerebral artery (STA-MCA) bypass + encephalo-duro-myosinangiosis (EDMS). The experimental group (Exp. Grp) was treated with ABMSC. The cerebral vascular tissue of the patients was treated with hematoxylin-eosin (HE) staining. Immunohistochemical staining was used to identify the levels of Cx43 protein. The concentrations of vascular endothelial growth factor (VEGF), inflammatory factor interleukin-6 (IL6), interleukin-1β (IL1β), tumor necrosis factor (TNFα), and anti-inflammatory factor interleukin-1β (IL1β) were determined by enzyme-linked immunosorbent assay (ELISA). We have found that after treatment of the expression of Cx43 protein, the proportions of grade IV (7.7%), grade III (311.5%), and grade II (3.8%) patients in the Exp. Grp were lower than those in the CG. The proportion of grade I patients in the Exp. Grp (77%) was higher than that in the CG (38.5%). After treatment, the inflammatory factors IL6 (0.97 ± 0.82 pg/mL), IL1β (8.33 ± 1.21 pg/mL), and TNFα (1.73 ± 0.71 pg/mL) in the Exp. Grp were lower than those in the CG. The anti-inflammatory factor IL1β (15.09 ± 4.72 pg/mL) increased in the Exp. Grp compared with the CG (11.25 ± 3.48 pg/mL) post treatment. Intracranial infection, hydrocephalus, hemiplegia, and transient neurological dysfunction in the Exp. Grp were lower than those in the CG, with statistical differences ($P < 0.05$). Our study suggests that the treatment of autologous bone marrow stem cells (ABMSC) was beneficial to balance the inflammatory response of disorders, reduce the damage of vascular tissue in the brain, and regulate tissue repair by co-acting with various inflammatory factors as compared to traditional surgery. We conclude that the involvement of Cx43 in the occurrence and development of moyamoya. We also have found that the risk factors of intracranial infection after ABMSCT were less as compared to those after conventional surgery.

1. Introduction

Moyamoya is a disease that comes in the rare category. It is related to the brain and its blood vessels with an unknown cause and a high mortality and disability rate. The incidence is about 4%, of which about 1.5% all have a family history [1, 2]. The main symptoms are narrowing or occlusion of bilateral internal arteries in the brain. Patients often have headaches, dizziness, epilepsy, aphasia, dyskinesia, and other uncomfortable symptoms, and the symptoms of some

patients are not particularly obvious [3, 4]. With the improvement in modern medical levels, moyamoya can be treated early and gradually. However, due to the defects of repeated attacks and the long duration of moyamoya, the occurrence may cause patients to suffer from temporary or permanent neurological damage, cognitive impairment, and other major diseases, and in more severe cases, death [5, 6]. It seriously affects the daily life and economy of patients and their families, making them suffer from psychological torture. Therefore, early detection and effective treatment can greatly improve the prognosis of patients and it is of great significance.

In Reference [7], the authors present a medical evaluation of MMS and moyamoya. They assess the epidemiological, pathological, and historical background of moyamoya They also present the clinical and radiographical outcomes and investigative imaging modalities. They highlight the efficacy of medical and surgical treatment and discuss the problems associated with surgical treatment. In Reference [8], the authors focus on the vasculopathy related to MYH11 mutations. They also state the effect of MYH11 mutations on cerebral arteriopathy. It is recommended that MYH11 testing could be thought of in the case of children suffering from moyamoya. In Reference [1], the authors state that the bone-marrow stem mobilization after revascularization in sufferers of moyamoya can make the recovery process faster and encourage the development of new blood vessels. It is able to lessen infection and improve patients' overall quality of life. In Reference [9], the authors study the clinical performance of multipoint cranium drilling and the usage of simvastatin as the remedy for moyamoya. Treating moyamoya by this method is safe and efficient, as it may assist in healing of nervous functions and it also improves patients' everyday life skills and quality of lifestyle. In Reference [10], the authors investigate if moyamoya can be caused due to mechanical stress caused by the blood flow that opposes the vulnerable intracranial vasculature. They also additionally look at the possibility that the angle between the supraclionoid segments and cavernous can be the cause of the disorder. The moyamoya ailment may result from mechanical factors acting at the ICA bifurcation. According to the research study, the angle may get increased because of blood vessel wall factors. In Reference [11], the authors point out the efficacy of surgical revascularization in heading off stroke in moyamoya. Surgical revascularization consists of augmenting the intracranial blood flow with the use of an external carotid system by use of direct bypass. The pial synangiosis can also be used. In Reference [12], the authors state that the moyamoya disease is found more amongst the younger patients. The patients are vulnerable to perioperative ischemic complications. They also state that the recognition of moyamoya might also contribute to the better surgical results obtained through perioperative management based on suitable surgical risk stratification. In References [13, 14], the authors try to identify whether or not extracranial-intracranial bypass can enhance patient prognosis and lessen the prevalence of rebleeding. They conclude that direct bypass lessens the danger of frequently

happening hemorrhages in moyamoya. In Reference [15], the Kaplan–Meier analysis is performed to reveal the difference between the nonsurgical and surgical groups. It also suggested the preventive effect of bypass on rebleeding.

In this research study, we look into the results of autologous bone-marrow stem cell therapy (ABMST) on the levels of inflammatory factors and Conexin43 (Cx43) protein in patients suffering with moyamoya. This research will serve as a guideline for the treatment of the disease.

The highlights of the research are given below:

(1) Identification of the population for the study

(2) Proposing the ABMSC treatment for moyamoya

(3) Determining the concentrations of VEGF, inflammatory factors IL6, IL1β, TNFα, and anti-inflammatory factor IL1β in the serum

(4) Cx43 immunohistochemical determination

(5) Statistical analysis is performed using the χ^2 test. For statistical processing, the t-test was used

The next sections of the paper elaborate the proposed work and results.

2. Materials and Methodology

The flow of the methodology is as follows:

(1) *Information gathering*: we gathered the information from the selected group. The inclusion and exclusion criteria have been mentioned in the section given below.

(2) *The treatment*: the CG (control group) was treated on the basis of conventional drug therapy, and the Exp. Grp was treated with ABMSC.

(3) *Staining*: HE staining and Immunohistochemical staining was performed in order to observe the cell level detailing.

(4) *ELISA*: It was used for the binding of antibodies to antigens. Cerebral vascular tissue samples were dissolved to determine the concentrations of VEGF, inflammatory factors IL6, IL1β, TNFα, and anti-inflammatory factor IL1β in the serum.

(5) *Cx43 immunohistochemical positive determination*: Immunohistochemical staining is utilized to find out the levels of Cx43 protein.

(6) *Analytical study*: In this study, SPSS13.0 was used for the analysis.

2.1. Basic Information and Grouping. 52 cases of patients from our hospital from 30 July 2019 to 10 February 2020 were selected, including 24 males, 28 females, 8 children, and 44 adults, with an average age of 42.56 ± 11.78. Among them, 9 patients had a history of hypertension and heart disease; 14 presented with intracerebral hemorrhage; and 38 presented with ischemic symptoms. Suzuki staging: 29 cases in phase 3, 23 cases in phase 4.

2.1.1. Inclusion Criteria.
The inclusion criteria were as follows:

(i) Patients with ages ranging from 18–65 yrs

(ii) Patients diagnosed with moyamoya disease/syndrome, and all patients were confirmed by DSA angiography

(iii) Initial symptoms presented as cerebral ischemia symptoms, such as transient ischemic attacks (TIA), limb numbness, visual defect, intermittent headache, dizziness, asymptomatic, or stroke (older than 3 months)

(iv) The admission's NIHSS score was less than 2

(v) The results of ECG and pulmonary function testing showed normal

2.1.2. Exclusion Criteria.
The exclusion criteria were as follows:

(i) DSA showed that moyamoya was complicated with other cerebrovascular ailments such as arteriovenous malformation (AVM), aneurysm, and posterior-circulation vascular disease.

(ii) Patients had a previous history of cerebral parenchyma hemorrhage, subarachnoid, and intraventricular hemorrhage.

(iii) Patients suffering from metabolic and endocrine system disorders.

52 moyamoya patients were divided into the CG and the Exp. Grp, each with 26 cases. The CG received conventional intracerebral and extracerebral vascular bypass surgery, while the Exp. Grp received the ABMSC.

2.2. The Treatment.
In the CG, on the basis of conventional drug therapy, all patients underwent a combined revascularization procedure which included indirect flow augmentation by EDMS and direct bypass through the STA-MCA bypass as previously described elsewhere (Peter book). Prior to performing the anastomosis, all patients underwent indocyanine green (ICG) and FLOW800 analysis to ensure the anastomosis side was in a low perfusion area in correlation with the radiological findings [16, 17]. We decided to perform STA-MCA bypass in relation to the hypoperfusion extension. The blood vessels cut from the anastomotic site were used as samples. The Exp. Grp was treated with ABMSC on the basis of the CG. After surgery, $2.5 \mu g/kg$ recombinant human granulocyte colony and granulocyte-macrophage colony-stimulating factors were alternately injected subcutaneously every three days for 21 consecutive days.

2.3. HE Staining.
The frozen cerebral vascular tissue was fixed with a 10% solution of formalin. Gradient dehydration was carried out with ethanol from a low to a high degree. Paraffin embedding was performed for $4 \mu m$ sections. The sections were routinely stained with HE, dewaxed with xylene, washed with ethanol at all levels, and then the staining was done for 10 min using hematoxylin and then rinsed out with tap water. After it was stained in eosin for 2 min, conventional ethanol was dehydrated, xylene was transparently sealed and fixed with neutral resin, and histopathological changes were examined by light microscopy. The nuclei were dark blue, and the cytoplasm and fibrous tissues were red in varying shades. The sections with better staining were photographed and analyzed.

2.4. Immunohistochemical Staining.
Immunohistochemical staining is a widely used technique used to identify the antigens present in biological tissue samples [18]. The vascular tissue in the brain was removed, and the sections were embedded by dehydration. Then, immunohistochemical staining was performed. The section was dehydrated and hydrated first. Antigen repair was performed to block endogenous peroxidase. It was washed with phosphoric acid buffer (PBS), and then goat serum blocking solution was added. Diluted primary antibody and biotinylated secondary antibody against goat anti-rabbit IgG antibody were added. The slides were washed for three times for 5 mins each, using PBS. Then, diaminobenzidine (DAB) was added for color development, the nuclei were re-stained with hematoxylin for 4 min, the cells were dehydrated with gradient ethanol, and the tablets were sealed with xylene transparent neutral resin. It was shot under a microscope to observe the vascular smooth muscle cell cytoplasmic expression of α-SMA. The positive color was tan, and the nuclei were re-stained light blue. The results of Cx43 immunohistochemical staining showed that the tan color was positive.

2.5. Enzyme-Linked Immunosorbent Assay (ELISA).
ELISA is used to measure antibodies, antigens, proteins, and glycoproteins in biological samples [19]. Cerebral vascular tissue samples were dissolved to determine the concentrations of VEGF, inflammatory factors IL6, IL1β, TNFα, and anti-inflammatory factor IL1β in the serum. The samples were incubated with the standard in the reaction well, and then a biotin-labeled primary antibody was added and it was incubated at room temperature for 2 h. After the reaction, the product was washed with buffer solution many times. The color solution $3,3',5,5'$-tetramethylbenzidine (TMB) was then added to each of the wells and it was incubated (at room temperature) away from the light. After the termination solution was added, the value of optical density (OD) was determined by an enzyme marker at a wavelength of 540 nm. The sample concentration was identified in accordance with the standard curve of the sample, and the actual concentration of the sample was obtained by multiplying the dilution factor.

2.6. Cx43 Immunohistochemical Positive Determination.
Conexin43 (Cx43) is the intercellular gap junction protein. It is responsible for the normal function of arteries. It also plays an important role in the development of cardiovascular

ailments. According to the positive cell percentage in the immunohistochemical results in the same kind of cells and the staining intensity of positive cells, the section results can be classified into grades I–IV by semiquantitative scoring. At the IV level, the positive cells accounted for more than 50%, and the positive cells were tan in color. At the III level, the positive cell percentage was 25%–50%, and the positive cells were brownish. At the II level, the positive cell percentage was 10%–25%, and the positive cells were yellowish in color. At the I level, the positive cell's percentage was less than 10%, and they were not stained.

2.7. Statistical Analysis. In this study, SPSS13.0 was used to analyze and process all experimental data. The counting data was represented by a percentage (%). A χ^2 test was performed. The data were expressed as mean ± standard deviation ($X \pm s$), and the t test was used for statistical processing. $P < 0.05$ indicated the differences that were statistically significant.

3. Results

3.1. Comparison of Age of Study Subjects. The mean age of the patients in the CG and the Exp. Grp and the ages of less than 19, 20–30, 31–40, 41–50, and more than 51 were compared and analyzed, as shown in Figure 1. The average age of the CG was 33.11 ± 10.02, and that of the Exp. Grp was 28.76 ± 11.8. The proportions of patients in the CG and the Exp. Grp aged less than 19, 20–30, 31–40, 41–50, and over 51 were (7.7%, 23.1%, 26.9%, 38.5%, and 3.8%) and (11.5%, 19.2%, 15.4%, 42.3%, and 11.5%), respectively. A statistical difference (stat. difference) was not found in the mean age and proportion of each age group between the CG and the Exp. Grp ($P > 0.05$). The following figure (Figure 1) shows the comparison of the average age of the patients and the proportion of the population in each age group in the CG and the experimental control.

3.2. MMD Diagnosed by DSA (Digital Subtraction Angiography). A DSA was performed on two moyamoya patients, and the results are shown in Figure 2. Figure 2(a) shows a 9-year-old child, female. The DSA image of the child showed bilateral middle cerebral arteries, occlusion of the anterior cerebral artery, and multiple tortuous collateral circulation vessels inside the skull. Figure 2(b) shows the DSA image of a 35-years-old, male adult suffering from moyamoya who showed a sudden cerebral infarction. The DSA showed occlusion of the middle cerebral artery and left internal carotid artery. It also shows the multiple localized stenosis of the right middle cerebral artery. The following figure (Figure 2) shows the DSA images obtained in two moyamoya patients.

3.3. HE Staining of Cerebral Blood Vessels in Moyamoya Patients. HE staining of cerebral blood vessels in moyamoya patients in the Exp. Grp and the CG are shown in Figure 3. The inner and middle membranes were thickened to different degrees. Figure 3(a) depicts severe thickening of the
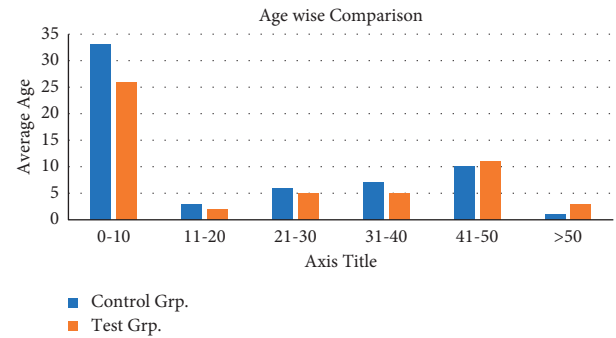
Figure 1: Comparison of the average age of the patients and the proportion of the population in each age group in the CG and the experimental control.

vascular intima with cells passing through the elastic membrane rupture. Figure 3(a) depicts moderate thickening of the intima, thinning of the medium-membrane, and interconnections between the intima and medium-membrane cells. Figure 3 shows the HE staining of cerebral blood vessels in moyamoya patients.

3.4. Cx43 Protein Expression and Immunohistochemical Staining Results. Immunohistochemical results of the Cx43 protein expression in the CG and the Exp. Grp after treatment are shown in Figure 4. In the CG, Cx43 protein in the cerebral blood vessels was mainly strongly positive or positively expressed, and the color was dark brown. Cx43 protein was not expressed or was only weakly positive in the Exp. Grp, and the color was light yellow. Figure 4 shows the immunohistochemical staining of the Cx43 protein expression.

The expression of Cx43 protein in the CG (26 cases) and the Exp. Grp (26 cases) after treatment were analyzed as shown in Figure 5. The Cx43 positive grading index showed that the proportion of grade IV patients (2 cases were 7.7%), grade III patients (3 cases were 11.5%), and grade II patients (1 case was 3.8%) in the Exp. Grp was lower than that of grade IV patients (6 cases were 23.1%), grade III patients (7 cases were 26.9%), and grade II patients (3 cases were 11.5%) in the CG. The proportion of grade I patients in the Exp. Grp (77% in 20 cases) was higher than that in the CG (38.5% in 10 cases), and there was a statistical difference between the two groups ($P < 0.05$). Figure 5 shows the Cx43 protein expression in CG and Exp. Grp after treatment.

3.5. ELISA Results of VEGF in CG (Control Group) and Exp. Grp. The results of VEGF ELISA before and after treatment in the CG and the Exp. Grp were compared and analyzed as shown in Figure 6. There was no significant difference in VEGF (3.44 ± 1.97 pg/mL) between the Exp. Grp and the CG (3.57 ± 2.09 pg/mL) ($P > 0.05$) before treatment. After treatment, the Exp. Grp (2.46 ± 2.11 pg/mL) was less than the CG (3.01 ± 1.54 pg/mL), and there was a stat. difference ($P < 0.05$). Figure 6 shows the ELISA results' comparison of VEGF before and after treatment between the CG and the Exp. Grp.
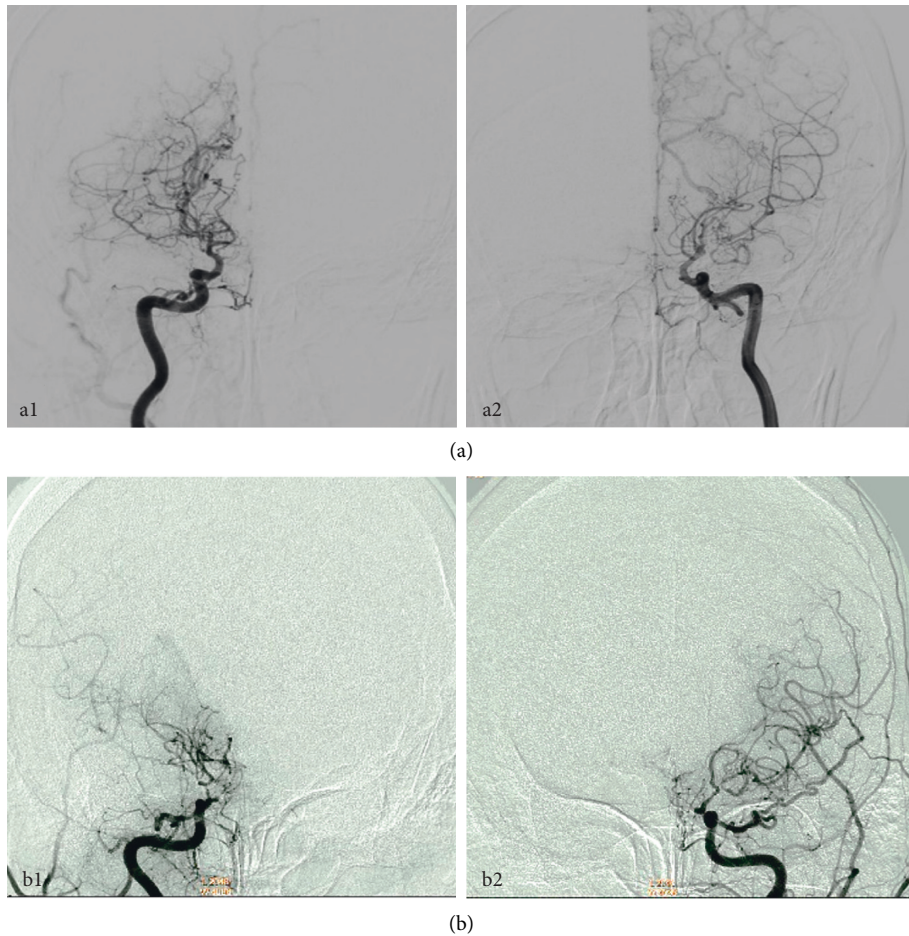
FIGURE 2: DSA images obtained in two moyamoya patients (a) a 9-year-old, female and (b) an adult, 35 years old male; the blue arrow pointed to the lesion area.
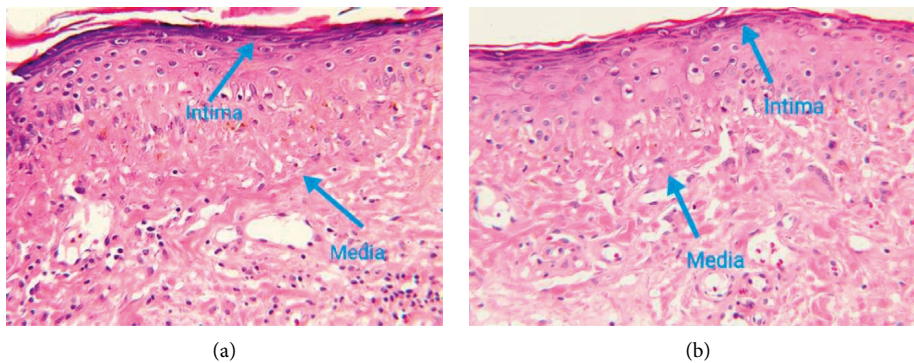
FIGURE 3: HE staining of cerebral blood vessels in moyamoya patients. (a) CG (control group) and (b) Exp. Grp. The blue arrows point to the inner and middle membranes of the vessels.

3.6. The ELISA Results Obtained from the CG and the Exp. Grp. The ELISA results of inflammatory factors IL6, IL1β, and TNFα were compared before and after treatment in the CG and the Exp. Grp. As shown in Figures 7(a)–7(c), the inflammatory factors IL6, IL1β, and TNFα in the Exp. Grp before treatment showed no ST as compared with the CG ($P > 0.05$). After treatment, the inflammatory factor IL6 (0.97 ± 0.82 pg/mL) in the Exp. Grp was lower than that in the CG (1.24 ± 0.45 pg/mL), and the inflammatory factor IL1β in the Exp. Grp (8.33 ± 1.21 pg/mL) was lower than that in the CG (10.13 ± 1.54 pg/mL), and the inflammatory factor TNFα (1.73 ± 0.71 pg/mL) in the Exp. Grp was lower than that in the CG (2.01 ± 1.03 pg/mL). After treatment, the inflammatory factors IL6, IL1β, and TNFα in the Exp. Grp were statistically (ST) different from those in the CG ($P < 0.05$).

(a)   (b)   (c)
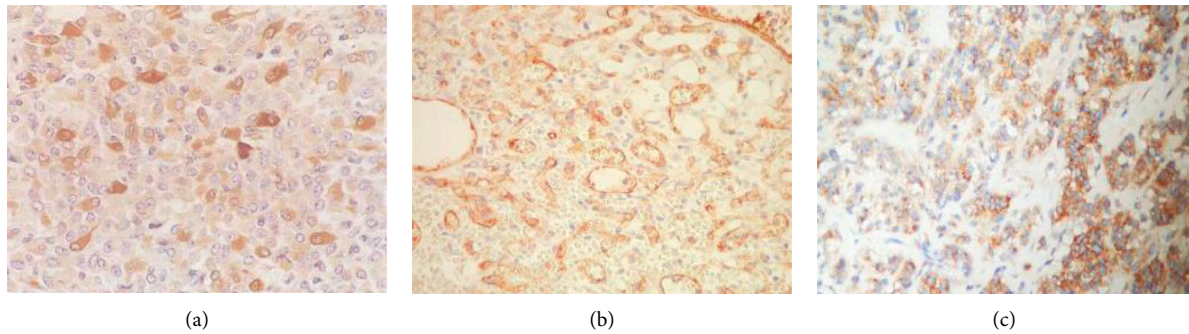
FIGURE 4: Immunohistochemical staining of the Cx43 protein expression. (a) Positive expression (Positive Exp.) of grade II Cx43 cells; (b) positive Exp. of grade III Cx43 cells; and (c) positive Exp. of grade IV Cx43 cells.
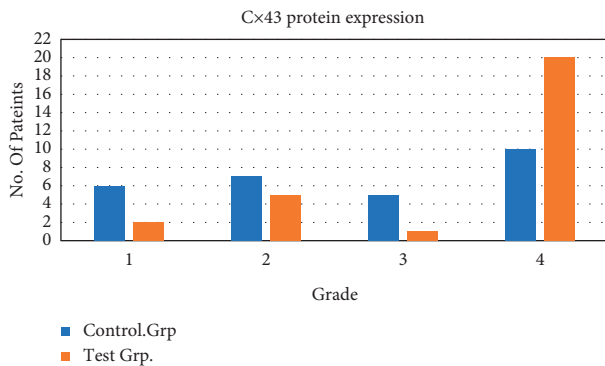
FIGURE 5: Cx43 protein expression in the CG and the Exp. Grp after treatment. *indicated that there was a statistical difference between the groups $P < 0.05$.
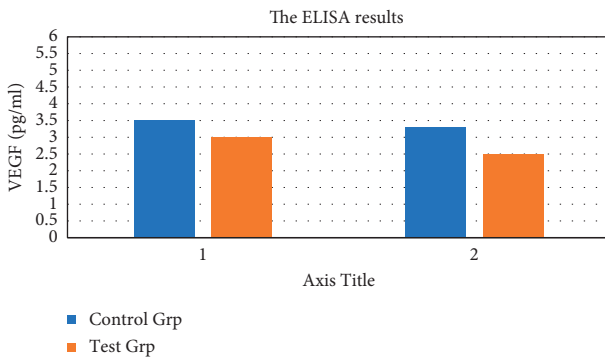
FIGURE 6: The ELISA results' comparison of VEGF before and after treatment between the CG and the Exp. Grp. *indicates that after treatment, there was a ST between the CG and the Exp. Grp, $P < 0.05$. (a) before treatment and (b) after treatment.

The ELISA results of the anti-inflammatory factor IL1β before and after treatment were compared between the CG and the Exp. Grp. As shown in Figure 7(d), there was no ST in the anti-inflammatory factor IL1β (5.58 ± 1.34 pg/mL) in the Exp. Grp compared with the CG (5.62 ± 1.01 pg/mL) before treatment ($P > 0.05$). After treatment, the anti-inflammatory factor IL1β (15.09 ± 4.72 pg/mL) increased in the Exp. Grp compared with the CG (11.25 ± 3.48 pg/mL), showing an ST with $P < 0.05$. The following figure (Figure 6) shows the ELISA results of related inflammatory factors in

the control group (CG) and the experimental group (EG) before and after treatment (A-IL6, B-IL1β, C-TNFα, and D-IL1β)

3.7. Complications after Treatment in the CG and the Exp. Grp. Complications such as intracranial infection, hydrocephalus, hemiplegia, and transient neurological dysfunction were analyzed in the CG and the Exp. Grp after treatment as shown in Figure 8. After treatment, intracranial infection (1 case 3.8%), hydrocephalus (0 cases), hemiplegia (0 cases), and transient neurological dysfunction (1 case 3.8%) were lower in the Exp. Grp than in the CG with intracranial infection (4 cases 15.4%), hydrocephalus (3 cases 11.5%), hemiplegia (2 cases 7.7%), and transient neurological dysfunction (1 case 3.8%). There was a ST between the two ($P < 0.05$).

4. Discussion

Moyamoya seriously affects the mental as well as the physical health of children and adolescents and is the main cause of cerebrovascular accidents in this population. In recent years, it has gradually attracted the attention of scholars from all over the world, but there is no effective method to control the occurrence and development of Moyamoya. In this study, 52 cases of moyamoya patients were selected for relevant research. Histopathological examination of the patient's intracranial blood vessels revealed rupture of the internal elastic membrane and uneven thickening of the vascular intima. It was found that there were endometrial cells and mesenchymal cells in the rupture of the internal elastic membrane, and the mesenchymal membrane was significantly thinner, which was consistent with mmd-related pathological changes and consistent with domestic and foreign literature. By comparing the expression of Cx43 protein in the CG and the Exp. Grp after treatment, the immunohistochemical results showed the proportions of grade IV patients (7.7%), grade III patients (311.5%), and grade II patients (3.8%) in the Exp. Grp were lower than that of grade IV patients (23.1%), grade III patients (26.9%), and grade II patients (11.5%) in the CG. The proportion of grade I patients in the Exp. Grp (77%) was higher than that in the CG (38.5%), and there was a statistical difference between

(a)

(b)

(c)

(d)
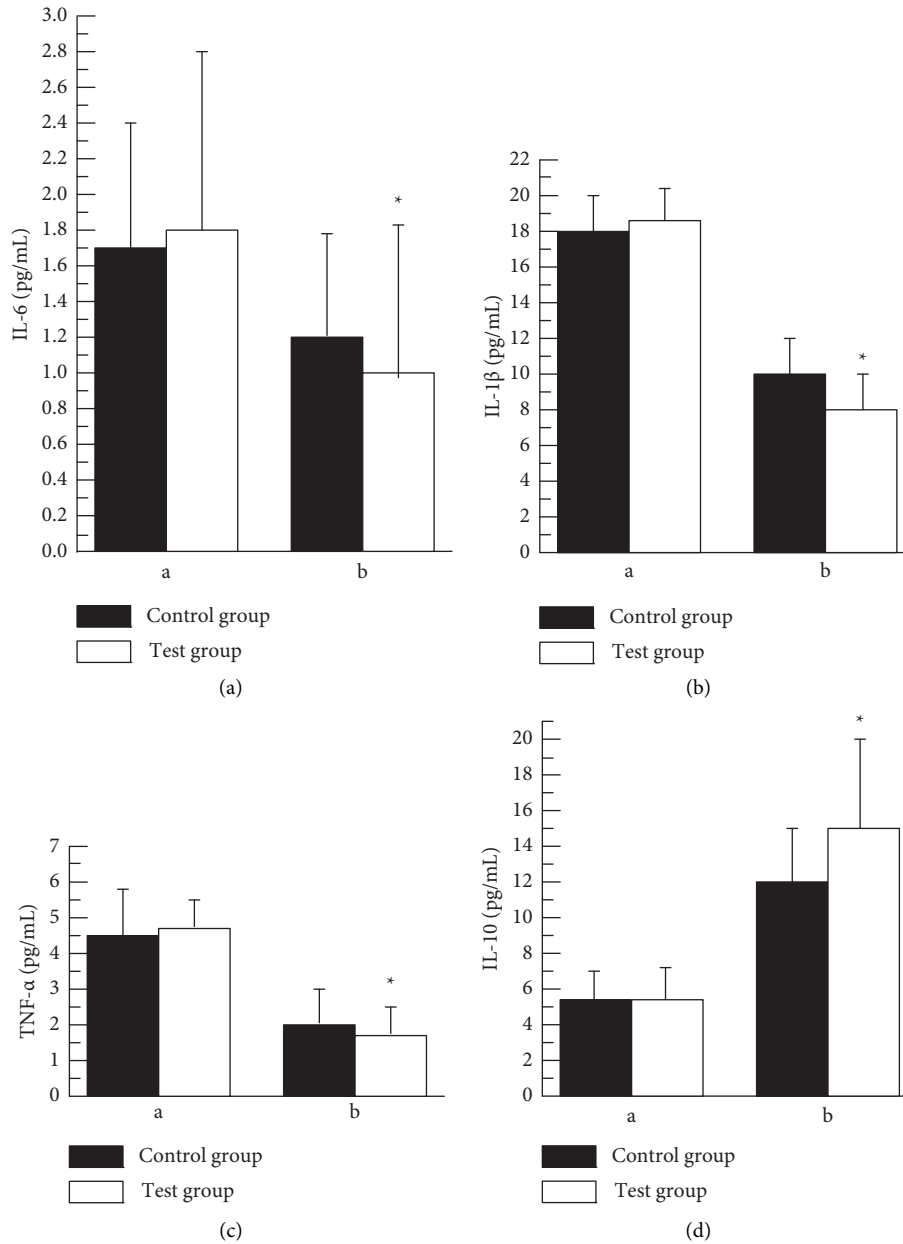
Figure 7: The ELISA results of related inflammatory factors in CG and the Exp. Grp before and after treatment ((a) IL6, (b) IL1β, (c) TNFα, and (d) IL1β). *indicates that after treatment, there was a ST between the CG and the Exp. Grp, $P < 0.05$ (a) before treatment and (b) after treatment.

the experimented groups ($P < 0.05$). Cx43 protein was not expressed or weakly positive in the middle and inner membranes of the blood vessels in the Exp. Grp, while Cx43 protein was strongly or positively expressed in the middle and inner membranes of the blood vessels in the CG. It was speculated that the Cx43 protein was involved in the migration as well as the proliferation of moyamoya in smooth muscle cells. VSMC or phenotypic transformation of VSMC; therefore, the Cx43 protein may be the cause of the development of moyamoya.

In this study, ELISA results of VEGF, inflammatory factors IL6, IL1β, TNFα, and anti-inflammatory factor IL1β were compared before and after treatment in the CG and the

Exp. Grp. After treatment, the Exp. Grp (2.46 ± 2.11 pg/mL) was lower than the CG (3.01 ± 1.54 pg/mL), showing a statistical difference (ST) with $P < 0.05$. In the pathological state, VEGF stimulated the formation and growth of tumor blood vessels and caused tissue edema in the inflammatory state. The increase of VEGF was also related to vascular diseases and played a certain role in the formation of moyamoya reticular blood vessels, which was in tandem with the results of Jeon [18]. After treatment, the inflammatory factors IL6 (0.97 ± 0.82 pg/mL), IL1β (8.33 ± 1.21 pg/ mL), and TNFα (1.73 ± 0.71 pg/mL) in the Exp. Grp were lower than those in the CG. After treatment, the anti-inflammatory factor IL1β (15.09 ± 4.72 pg/mL) in the Exp. Grp
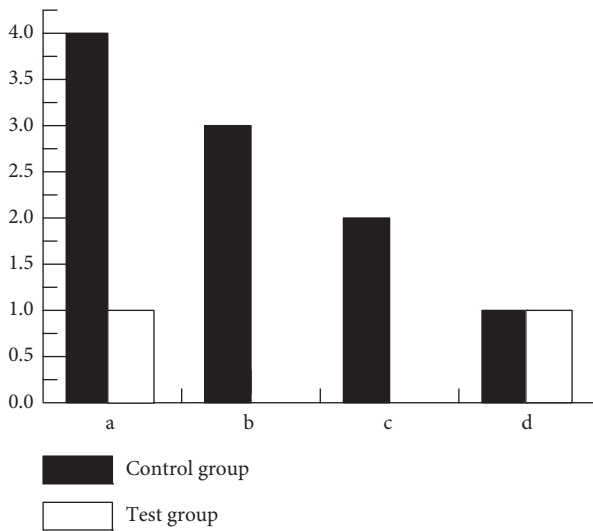
FIGURE 8: Complications of (a) intracranial infection, (b) hydrocephalus, (c) hemiplegia, and (d) transient neurological dysfunction in the CG and the Exp. Grp after treatment. *indicates that there was a ST between the CG and the Exp. Grp, $P < 0.05$.

increased compared with that in the CG(11.25 ± 3.48 pg/mL), and the difference was statistically significant ($P < 0.05$). After ABMSCT in moyamoya patients, the inflammatory factors IL6, IL1β, and TNFα in plasma can be reduced, and the anti-inflammatory factor IL1β can be increased. It indicated that ABMSC has an intervention effect on the secretion of various inflammatory factors of moyamoya, thus providing a repairable microenvironment for the vascular tissues in the brain and playing a protective role on the tissues [17].

The analysis of the prognostic complications showed that the intracranial infection, hydrocephalus, hemiplegia, and transient neurological dysfunction after treatment in the Exp. Grp were lower than those in the CG, and there was a statistical difference between the two ($P < 0.05$), which indicated that the prognosis of ABMSC for moyamoya was better, and the physiological indicators of the patients have been significantly improved, and the adverse reactions of the patients have not been increased.

5. Conclusion

In our study, the effects of ABMSCT on the levels of inflammatory factors and Cx43 protein in patients with moyamoya have been analyzed. The results show the basic pathological changes of moyamoya patients resulted in an uneven thickening of the cerebral vascular intima. It is concluded that the involvement of Cx43 in the occurrence and development of moyamoya is significant. It can also be summarized that the treatment of ABMSC is beneficial to balance the inflammatory response of the disorder, reduce the damage of vascular tissue in the brain, and regulate the tissue repair by co-acting with various inflammatory factors, which provided certain guidelines for the clinical treatment of moyamoya. It also investigated the risk factors of intracranial infection after autologous bone-marrow stem cell

therapy. It has been found that the rates of intracranial infection, hydrocephalus, hemiplegia, and transient neurological dysfunction after treatment in the patients treated with ABMSC are lower than in those patients who were treated on the basis of conventional drug therapy. Our future progress will include follow-up analysis of the prognosis of patients after autologous bone marrow stem cell treatment, which can be further confirmed by the data.

Data Availability

The data can be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest with respect to this article.

References

[1] L. Zhao, W. Sun, H. Liang et al., "Therapeutic effect of autologous bone marrow stem cell mobilization combined with anti-infective therapy on moyamoya disease," *Saudi Journal of Biological Sciences*, vol. 27, no. 2, pp. 676–681, 2020.

[2] J. Liao, T. Hong, J. Xu, E. Zeng, B. Tang, and W. Lai, "Expression of Connexin43 in cerebral arteries of patients with moyamoya disease," *Journal of Stroke and Cerebrovascular Diseases*, vol. 27, no. 4, pp. 1107–1114, 2018 Apr.

[3] I. Yamaguchi, J. Satomi, N. Yamamoto et al., "Coexistence of quasi-moyamoya disease and poems syndrome in a patient with intracranial hemorrhage: a case report and literature review," *NMC Case Report Journal*, vol. 4, no. 1, pp. 5–9, 2017.

[4] G. Bedini, K. Blecharz, S. Nava et al., "Vasculogenic and angiogenic pathways in moyamoya disease," *Current Medicinal Chemistry*, vol. 23, no. 4, pp. 315–345, 2016.

[5] J. P. Jeon, J. E. Kim, W. S. Cho, J. S. Bang, Y. J. Son, and C. W. Oh, "Meta-analysis of the surgical outcomes of symptomatic moyamoya disease in adults," *Journal of Neurosurgery*, vol. 128, no. 3, pp. 793–799, 2018 Mar.

[6] Y. C. Wei, C.-H. Liu, T.-Y. Chang et al., "Coexisting diseases of moyamoya vasculopathy," *Journal of Stroke and Cerebrovascular Diseases*, vol. 23, no. 6, pp. 1344–1350, 2014.

[7] J. A. Berry, V. Cortez, H. Toor, H. Saini, and J. Siddiqi, "Moyamoya: an update and review," *Cureus*, vol. 12, no. 10, Article ID e10994, 2020 Oct 16.

[8] A. Keylock, Y. Hong, D. Saunders et al., "Moyamoya-like cerebrovascular disease in a child with a novel mutation in myosin heavy chain 11," *Neurology*, vol. 90, no. 3, pp. 136–138, 2018.

[9] N. Su, Z. Zhang, F. Jia, and R. Wu, "Treatment of moyamoya disease by multipoint skull drilling for indirect revascularization combined with mobilization of autologous bone marrow stem cells," *Genetics and Molecular Research*, vol. 14, no. 3, pp. 7519–7528, 2015.

[10] B. J. Sudhir and S. K. Kumar, "Moyamoya disease: using computational fluid dynamics to propose a novel "mechanical theory" of pathogenesis," in *Proceedings of the 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, pp. 72–77, IEEE, Bangkok, Thailand, October 2018.

[11] T. Kim, C. W. Oh, J. S. Bang, J. Kim, and W.-S. Cho, "Moyamoya disease: treatment and outcomes," *Journal of stroke*, vol. 18, no. 1, pp. 21–30, 2016.

[12] M. Kaur, S. R. Sakhare, K. Wanjale, and F. Akter, "Early stroke prediction methods for prevention of strokes," *Behavioural Neurology*, vol. 2022, Article ID 7725597, 9 pages, 2022.

[13] S. Miyamoto, T. Yoshimoto, N. Hashimoto et al., "Effects of extracranial-intracranial bypass for patients with hemorrhagic moyamoya disease: results of the Japan Adult Moyamoya Trial," *Stroke*, vol. 45, no. 5, pp. 1415–1421, 2014 May.

[14] N. M. Mule, D. D. Patil, and M. Kaur, "A comprehensive survey on investigation techniques of exhaled breath (EB) for diagnosis of diseases in human body," *Informatics in Medicine Unlocked*, vol. 26, Article ID 100715, 2021.

[15] Y. Lei, Y. Li, L. Yu et al., "Faded critical dynamics in adult moyamoya disease revealed by EEG and fMRI," *Oxidative Medicine and Cellular Longevity*, vol. 2021, Article ID 6640108, 17 pages, 2021.

[16] J. Chan, D. F. Ambrosio Rodriguez, D. Shalini, and C. Boucher-Berry, "Moyamoya disease with coexistent hypertriglyceridemia in pediatric patient," *Case Reports in Endocrinology*, vol. 2016, Article ID 7974182, 5 pages, 2016.

[17] W. Miao, P.-L. Zhao, Y.-S. Zhang et al., "Epidemiological and clinical features of Moyamoya disease in Nanjing, China," *Clinical Neurology and Neurosurgery*, vol. 112, no. 3, pp. 199–203, 2010.

[18] M. Jeon, "Emotions and affect in human factors and human-computer interaction: taxonomy, theories, approaches, and methods," in *Emotions and Affect in Human Factors and Human-Computer Interaction*, pp. 3–26, Academic Press, Cambridge, 2017.

[19] B. Mareschal, M. Kaur, V. Kharat, and S. Sakhare, "Convergence of smart technologies for digital transformation," *Tehnički glasnik*, vol. 15, p. 1, 2021.

Retraction

Retracted: Neural Network Based on Health Monitoring Electrical Equipment Fault and Biomedical Diagnosis

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. Zhang and Y. Lyu, "Neural Network Based on Health Monitoring Electrical Equipment Fault and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8358794, 7 pages, 2022.

Research Article

Neural Network Based on Health Monitoring Electrical Equipment Fault and Biomedical Diagnosis

Xinjun Zhang and **Yingli Lyu**

Department of Electrical Engineering, Jiyuan Vocational and Technical College, Jiyuan 459000, China

Correspondence should be addressed to Xinjun Zhang; 19404386@masu.edu.cn

In order to improve the accuracy of electrical equipment failure diagnosis and keep electrical equipment operating safely and efficiently, this paper proposes to design an electrical equipment failure diagnosis system based on a neural network, analyze the faults of electrical equipment and their causes, and establish knowledge base according to relevant data and expert judgment. The fault knowledge base was introduced into the neural network operation structure, and the fault diagnosis results were classified step by step through multiple subnetworks. In data preprocessing, in order to avoid the redundancy of primary fault information features, the principal component heuristic attribute reduction algorithm was used to select the fault data samples optimally. The neural network learning algorithm is used to calculate the forward direction and error rate of the initial error data, and the reliability function is used to optimize the initial weight threshold of the neural network, propagating the error backwards and high. Experimental results show that adding attribute reduction improves error classification performance, avoids the problem of local minima through neural network operation, and has fewer iteration steps, lower average error, and higher accuracy of fault diagnosis, reaching 95.6%.

1. Introduction

In power system, electrical equipment is an important core to ensure power operation and transmission. In actual operation, the fault ratio of electrical equipment is high [1]. At present, the fault inspection and maintenance of electrical equipment mostly depend on the experience and judgment of the staff. It can no longer meet the operational requirements of modern energy systems [2]. Therefore, it is very important in practice to continue working on failure diagnosis technology for electrical equipment. In recent years, some scientists have been studying in this direction. Reference [3] proposes a framework for electrical equipment failure diagnosis systems based on MultiAgent technology. For monitoring individual electrical equipment, ODS is used to access Power Big Data which improves the access efficiency of Power Big Data. Management and control agent, analysis and diagnosis agent, and information-aware agent are used to realize the intelligentization of the fault diagnosis

system. Under the environment of cooperative working mechanism, the specific tasks of agents are determined. The multiagent system is established to realize intelligent diagnosis of multiple electrical devices, but this method does not consider the fault diagnosis associated with multiple electrical devices, which has certain limitations. In reference [4], it is proposed to analyze the dissolved gas content in oil in electrical equipment by gas chromatograph to judge the hidden faults in electrical equipment. Taking a transformer as an example, the acetylene content in the oil chromatographic data of the pressure side oil tank exceeds the standard, and the opening inspection of the pressure side oil tank shows that the last screen lead-out line of the connecting flange is loose, resulting in discharge, which shows the effectiveness of the oil chromatographic data analysis method. However, this method is only suitable for the hidden faults in oil-filled electrical equipment, and its scope of application is limited. In reference [5], a hierarchical Bayesian fault diagnosis method based on kernel is

proposed. The kernel method is introduced into the hierarchical Bayesian model, optimized by kernel principal component analysis, and a calculation model with simplified parameters is obtained by expectation maximization algorithm. Multiclass predictive diagnosis of circuit breaker faults is realized by combining fault tree, but the calculation process of this method is complex and the calculation redundancy is high. There are other inspection methods in Japan and abroad. B. Wavelet-based information filtering for fault diagnosis of electric propulsion systems on electric ships studies the faults of electric drives in electrical systems, filters rich information by using two-dimensional wavelet transform of sensor data, reduces the computational complexity of classifiers, and realizes fault diagnosis in electric drive sensors of electric ships [6]. There is also research on failure diagnosis based on electrical signature analysis for predictive maintenance of synchronous generators in large electrical systems. By applying the electrical characteristic analysis technology to the condition monitoring research of static var generator in large power system, the faults of stator electrical imbalance and mechanical dislocation are detected, and the electrical characteristics of wound rotor static synchronous generator are also discussed, including signal analysis suggestions, how to identify rotor faults according to fault modes, and the characteristics of bipolar static synchronous generator [7]. In addition, the traditional methods mentioned above are not standardized in obtaining fault data, and it is difficult to obtain useful information. Moreover, they rely on a comprehensive case base and expert experience, which brings great difficulties for accurate fault diagnosis.

Electrical equipment has complex structure and various fault modes, and the core of its diagnosis is fault classification [8]. The structure of neural network is simple, which is suitable for solving the problem of complex internal mechanism. It can approach any nonlinear function with arbitrary precision and adjust the threshold of network weights repeatedly through reverse learning, so as to minimize the difference between actual and expected output. Therefore, this paper proposes to design electrical equipment fault diagnosis system based on neural network. By consulting relevant literature and expert research data, the knowledge base of historical fault information diagnosis of electrical equipment is constructed to provide data support for the fault diagnosis system. The neural network architecture of electrical equipment fault diagnosis is constructed [9]. In order to improve the speed of fault diagnosis and classification, fault data are processed by multiple subneural networks, and data are transmitted in parallel and fused. In the neural network operation, the heuristic reduction algorithm of attribute principal component is used to construct the component function of attribute on the basis of difference matrix, which simplifies the original sample set and improves the training efficiency. Forward and error calculations of error data are performed by neural network. Based on this, the confidence function is used to optimize the initial weights and thresholds of the neural network, and the optimized results are captured and trained in the neural network. This provides reverse error transmission. It not only gives full play to the mapping ability of neural network generalization but also avoids the local minimum problem and improves the diagnosis accuracy. The effectiveness of this method in fault diagnosis of electrical equipment is verified by experiments.

2. Electrical Equipment Fault Diagnosis Knowledge Base Construction

The operation of electrical equipment needs the cooperation of various electrical components, and the fault is usually caused by the loss of control of electric energy or control information in the process of transmission, distribution, and conversion [10]. Common faults of electrical equipment include open circuit, short circuit, abnormal grounding, electric leakage, damage of electrical components, output error of electronic equipment due to electromagnetic interference, and accidental failure of control system components [11]. Electrical equipment failure may lead to extensive loss of personnel and property. Therefore, it is of great significance for fault diagnosis to collect the phenomena and causes of electrical equipment faults with high probability. The electrical equipment fault diagnosis knowledge base is constructed, and the electrical equipment fault characterization and fault causes are collected which includes basic facts, rules, and other relevant information [12, 13]. The information in the knowledge base comes from consulting a large number of literature studies and many domain experts, which is the key to determine the inference ability of the system. Some diagnostic knowledge is shown in Table 1.

In the system, the expression form of knowledge base is production rules, each rule corresponds to a conclusion, and the premises are represented by "AND". For example, IF winding is overheated AND fuse is blown, THEN coil is heated.

3. Design of Fault Diagnosis Structure of Electrical Equipment Based on Neural Network

Neural network is a mathematical model derived from the human nervous system, which is an algorithm model for distributed parallel information processing [14]. This model processes information through a large number of neurons with nonlinear mapping ability. In the network, neurons are organized in the form of hierarchical structure [15–18], and the processing units on each layer are connected with neurons on other layers in a weighted way. Using error back propagation, a three-layer forward neural network model is obtained, with the specific structure as shown in Figure 1.

By studying many typical fault samples in the knowledge base, we can recall the flaws' associated characteristics. When a case is entered, the neural network will use associative memory to find the closest fault and then realize fault diagnosis. However, when diagnosing multiple fault modes, the neural network will have too many nodes, resulting in huge structure, redundant information, and slow processing speed [19]. Therefore, this paper designs and uses several

TABLE 1: Knowledge base for fault diagnosis of electrical equipment.

Fault symptom	Cause of failure
The two-phase insulation resistance of the motor is too low, the insulation of the wire skin is damaged, the coil is hot, and the vibration sound of the unit is loud	Winding overheating, winding insulation breakdown, fuse burning, and improper stator offline
The insulation resistance between stator winding conductor and iron core is too low, and the motor is overheated	The end of insulation aging winding touches the end cover, the power supply voltage is three-phase asymmetric, the voltage is too high, and the stator winding is short-circuited
Active and reactive loads of the unit decrease	Transmitter misoperation and feedback sensor failure
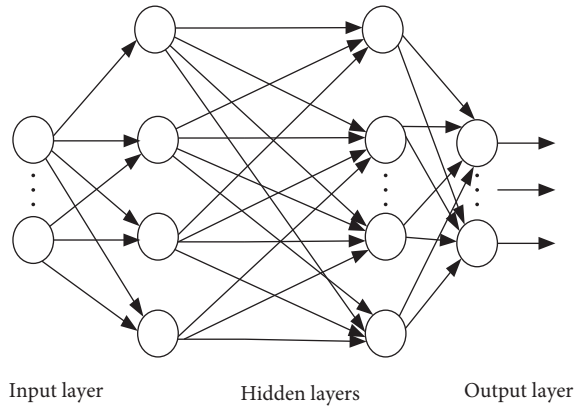
FIGURE 1: Structure diagram of neural network.

FIGURE 2: Structure diagram of neural network for fault diagnosis of electrical equipment.

subneural networks to diagnose faults from different sides, and then fuses the diagnosis results of the subnetworks, thus reducing the uncertainty of fault diagnosis and improving the recognition rate.

The neural network for fault diagnosis of electrical equipment designed in this paper is composed of three layers of networks connected in series [20, 21], which has the characteristics of fast data transmission rate in series structure. The specific structure is shown in Figure 2. In layer 2, each subnetwork is connected in parallel so as to avoid the phenomenon that information is not transmitted if a certain network fails in the series structure.

The fault diagnosis steps of electrical equipment mainly include following[22–24]:

(1) collecting fault information and converting it into an input mode of a neural network

(2) associating the fault knowledge base and outputting information classification results through forward calculation in each subneural network

(3) The classification results are transmitted to the decision fusion network, and the accurate fault diagnosis results are obtained by global fusion calculation

4. Fault Diagnosis Algorithm of Electrical Equipment Based on Neural Network

4.1. Data Preprocessing. The electrical equipment fault data samples collected by each sensor are transmitted to the neural network, and after being associated with the fault knowledge base [25, 26].

Because of the redundancy between data, the spatial vector model is established by selecting feature words, and the original data are processed by the heuristic attribute reduction algorithm of principal components [27, 28], and the optimal selection is made to complete the data preprocessing. First, we need to extract fault feature words from fault statements, and the calculation formula is as follows:

$$A = \frac{h \sum_{i=1}^{n} x_i t}{\delta}. \tag{1}$$

In the formula, A represents fault feature words, h represents fault feature word space vector, t represents correlation degree with fault knowledge base, δ represents fault feature word selection operator, and $x_i\,(i = 1, 2, \ldots, n)$ represents fitness at the i fault data acquisition node. Chi-square test theory is used to select features, which can effectively measure the correlation between feature items and categories. The calculation formula is as follows:

$$B = \frac{|t - p|^{\alpha}}{I - A\cot a}. \tag{2}$$

In the formula, p represents the fault probability of electrical equipment, α is the fitness population, I is the imaginary number of the fault itemset in the corresponding fault knowledge base, and a is the constant.

Through the word frequency-inverse file frequency, the corresponding weight is calculated based on the selected characteristic frequency [25], and the conversion from words to vector space is completed. The calculation formula is as follows:

$$C = T \sqrt{\frac{B - I \cot a}{\pi p^2}}. \tag{3}$$

In the formula, T represents the frequency of fault feature words in the text.

In this way, the primary selection of electrical equipment fault characteristics is completed. But at this time, redundancy and compatibility exist among fault features, which will make the calculation of neural network complicated [26]. The principal component heuristic attribute reduction algorithm is used for reduction. The attribute weights of initial features are discretized by clustering. The number of clusters is l, and the calculation formula is as follows:

$$D = \sum_{i=1}^{n} p w_i + \frac{c \sqrt{d}}{Cl}. \tag{4}$$

In the formula, w_i represents the component function of a certain attribute, c is the conditional attribute, and d is the decision attribute.

Through the principal component heuristic algorithm, the component values of each attribute are calculated by using the difference matrix, and the reduced fault characteristics are obtained. The calculation formula is as follows:

$$E = D \left\{ \frac{[\sqrt{f} - X]}{[cf + d]^2} \right\}. \tag{5}$$

In the formula, $X = (u_{ij})$ represents the difference matrix, u_{ij} represents the elements of the i row and j column in the difference matrix, and f represents the specific value of the fault data on the conditional attribute.

After reduction, the features of fault decision are greatly reduced, the number of input nodes of neural network is reduced, the structure of neural network is simplified, the redundancy between fault data is removed, and the operation cost is reduced.

4.2. Subneural Network Learning Operation. The number of neurons in the input layer, hidden layer, and output layer is L, M, and N. y_{LM} and y_{MN} are the thresholds for the input and hidden layers, and the hidden and output layers, respectively. The activation functions of hidden layer and output layer are $g(m)$ and $g(n)$, respectively. β and γ represent the input and output of neurons, respectively. The sample set of fault data is $s_k = (s_1, s_2, \ldots, s_k), k = 1, 2, \ldots, L$. The expected output is $d_k = (d_1, d_2, \ldots, d_k), k = 1, 2, \ldots, N$. The hidden layer output is as follows:

$$\gamma_M = E g(m) \sum_{k=1}^{L} s_k + \sum y_{LM}. \tag{6}$$

The output of the network is as follows:

$$\gamma_N = E g(n) \sum_{k=1}^{N} d_k + \sum y_{MN}. \tag{7}$$

The forward calculation of the subneural network is finished at this point.

The second neuron's output error in the output layer is as follows:

$$o_k = d_k - (\gamma_M - \gamma_N). \tag{8}$$

The sum of error energies of all neurons in the output layer is as follows:

$$W = \frac{1}{N} \sum_{k=1}^{N} o_k^2. \tag{9}$$

The reliability function is introduced into the subneural network structure to identify the possible faults in electrical equipment and the reliability of fault judgment according to certain judgment rules [27]. According to the output error obtained above, adjust the subneural network's weight threshold. The following formula is used to calculate the weighting factor between the output layer and the concealed layer:

$$K = |W|^2 - \frac{2}{\theta}. \tag{10}$$

In the formula, θ indicates credibility.

The adjusted weight threshold is transferred back to the input layer, and the weight coefficient between the input layer and the hidden layer is adjusted as follows:

$$Q = (\eta + \phi) K. \tag{11}$$

In the formula, η represents the numerical gain coefficient and ϕ represents the inertia coefficient. The learning convergence speed is adjusted by these two coefficients, which are usually between [0,1]. The fault feature of the subneural network is recalculated, and the reverse transmission of error is realized.

4.3. Neural Network Decision Fusion. After completing the fault feature output of the subneural network, the feature vectors are fused, combined with random disturbance vector $\varphi(k) = (\varphi_1(k), \varphi_2(k), \varphi_i(k))^T$, after normalization, and the calculation formula is as follows:

$$R = o_N \varphi_k \cos\left(\frac{2\pi}{W}\right) (\gamma_M + \gamma_N)^2. \tag{12}$$

By adding random disturbance value, enlarging the weight adjustment amount, speeding up the training speed of the network, and avoiding the parameter adjustment from entering the saturation state when the network parameters enter the saturation region, the neural network will jump out of the local minimum when fusing, ensuring that the network parameter adjustment will continue to keep the overall error convergence direction and complete the design of the failure diagnosis system for electrical equipment based on neural networks.

5. Experimental Results and Analysis

To test the application performance of the developed neural network-based electrical equipment malfunction diagnosis system, simulation experiments are needed. The neural

TABLE 2: Reliability results of electrical equipment fault diagnosis.

State code	Fault sample	Credibility
1	Failure 1	0.99
2	Failure 2	0.98
3	Failure 3	0.98
4	Failure 4	0.97
5	Failure 5	0.98
6	Failure 6	0.98
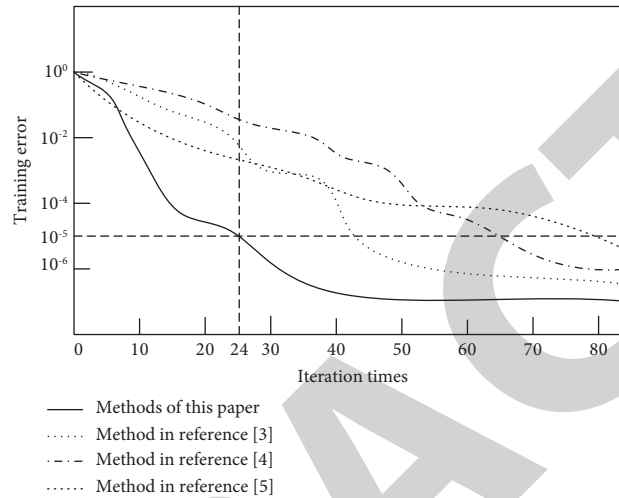
FIGURE 3: Comparison results of neural network error curves.

network is trained by MATLAB simulation platform, the training function is trainlm, the learning rate is 0.1, the target error is 10^{-5}, and the maximum number of iterations is 1000. The number of initial failure populations is 40, the selection parameter is 0.08, the crossover operator is 0.75, and the mutation probability is 0.01. Six kinds of fault voltage samples output by an electrical equipment are selected as test samples, and the reliability of fault diagnosis is tested by this method. The results are shown in Table 2.

It can be seen from the table that the reliability of this method in the fault diagnosis output of electrical equipment is higher, always above 0.97, which shows that the fault diagnosed is effective. The method of distributed electrical equipment fault diagnosis system based on multiagent technology in document [3], the application research method of oil chromatography analysis in oil-filled electrical equipment fault diagnosis in document [4], and the method of hierarchical Bayesian circuit breaker fault diagnosis based on kernel in document [5] is compared with the method in this paper. To comprehensively evaluate the performance of your design method, the number of convergence steps, the average absolute error of test samples, the maximum absolute error of test samples, and the accuracy rate are used to judge. The training error is shown in Figure 3.

From the figure, we can see that the method in this paper converged after 24 iterations, while the convergence iteration times of the methods in reference [3], reference [4], and reference [5] are 42, 65, and 80, respectively. By analyzing Figure 3, it can be seen that the method in this paper has

reduced the error quickly with fewer iterations, while the error reduction rate of other methods is slow and fluctuates greatly. This is because this method preprocesses the fault data by using the heuristic reduction algorithm of attribute principal component, which greatly reduces the redundancy of calculation and can get the fault data processing result by using fewer iterations. The average absolute error, maximum absolute error, and accuracy of test samples are shown in Table 3.

After adding attribute reduction algorithm, the average absolute error is 4.73%, the accuracy rate is 95.6%, the average absolute error is 6.84%, the accuracy rate is 88.6%, the average absolute error is 8.4%, the accuracy rate is 90.5%, and the average absolute error is 10.35%. The accuracy of this method is far superior to other methods, and both the average absolute error and the maximum absolute error are far less than other methods, which show that the attribute reduction algorithm weakens the compatibility between fault data and is effective in classifying the original data.

In order to verify the anti-interference performance of the system designed in this paper, under different interference signal-to-noise ratios, the fault diagnosis speed of this method is compared with that of literature [3], literature [4], and literature [5], and the comparison results are shown in Table 4.

According to Table 4, when the interference signal-to-noise ratio is 48 dB, the fault diagnosis rate of this method is the longest, which is 5.86 ms, and the shortest, which is 4.32 ms, when the interference signal-to-noise ratio is 56 dB.

TABLE 3: Comparison of experimental results.

Experimental methods	Average absolute error of test sample (%)	Maximum absolute error of test sample (%)	Accuracy (%)
Methods of this paper	4.73	9.24	95.6
Method in reference [3]	6.84	13.26	88.6
Method in reference [4]	8.4	16.38	90.5
Method in reference [5]	10.35	20.72	92.1

TABLE 4: Comparison of fault diagnosis rate.

Interference SNR/dB	Methods of this paper/ms	Method in reference [3]/ms	Method in reference [4]/ms	Method in reference [5]/ms
12	5.52	28.34	25.88	18.61
28	5.42	29.12	24.94	19.54
36	5.57	32.12	25.34	19.76
48	5.86	33.46	28.64	20.45
56	4.32	34.75	27.41	19.86
65	4.98	30.34	28.64	21.64
79	5.63	31.36	29.35	22.47

The method in reference [3] has the longest fault diagnosis rate of 34.75 ms and the lowest rate of 28.34 ms under the interference signal-to-noise ratio of 56 dB. The method in reference [4] has the longest fault diagnosis rate of 29.35 ms and the lowest rate of 24.94 ms under the interference signal-to-noise ratio of 79 dB. The method in reference [5] has the longest fault diagnosis rate of 22.47 ms and the lowest rate of 19.54 ms under the interference signal-to-noise ratio of 79 dB. Comparing the fault diagnosis time of the four methods, it can be seen that the electrical equipment fault diagnosis rate of this method is the highest and the diagnosis efficiency is the best, which shows that the electrical equipment fault diagnosis system based on the neural network designed in this paper has better performance and can realize the electrical equipment fault diagnosis effectively and accurately.

6. Conclusion

With the continuous development of industrial enterprises, electrical equipment has been widely used. However, due to its frequent failures and diversified causes, its diagnosis has become a difficult point. In order to diagnose faults in electrical equipment efficiently and accurately, this paper proposes to design a fault detection system for electrical equipment based on neural network. Based on the establishment of electrical equipment fault knowledge base, several subneural networks are used to transfer the fault data so as to reduce the operation time and improve the efficiency of fault diagnosis. The attribute reduction algorithm is used to preprocess the fault data, eliminate the redundant information in the fault data, reduce the dimension of the fault information, extract the main characteristic parameters, optimize the original fault data, and effectively improve the operation speed and fault classification ability. Considering the elimination of diagnostic errors and deviations, the reliability function is introduced to optimize the initial weights and thresholds of the neural network when the neural network is used for the forward calculation and error

calculation of fault data, and the reverse transmission training of errors is carried out to avoid the problem of local minima. Experiments show that the fault diagnosis system of electrical equipment based on neural network designed in this paper can effectively diagnose faults, with high reliability, less iteration steps, low average error, and fault diagnosis accuracy of 95.6%, which can provide theoretical support for fault diagnosis of electrical equipment and lay a foundation for practical application.

Data Availability

Data used to support this study are available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Cui, "Optimization of preventive maintenance cycle of ship mechanical and electrical equipment based on MRO system," *Journal of Coastal Research*, vol. 93, no. sp1, p. 953, 2019.

[2] N. Zhou, L. Luo, G. Sheng, and X. Jiang, "High accuracy insulation fault diagnosis method of power equipment based on power maximum likelihood estimation," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1291–1299, 2019.

[3] T. Yang, Z. Zeng, and L. Zhou, "Research on Distributed Electrical Equipment Fault Diagnosis System Based on Multi-Agent and Big Data," *Distribution & Utilization*, vol. 10, pp. 74–78, 2018.

[4] Z. Huang, P. Chen, and C. Dun, "Application of oil chromatographic analysis in fault diagnosis of oil-filled electrical equipment," *Rural Electrification*, vol. 02, pp. 19–22, 2019.

[5] H. Zhu and Y. Ding, "Kernel based hierarchical Bayesian fault diagnosis method for circuit breaker[J]," *Chinese Journal of Construction Machinery*, vol. 5, pp. 450–454, 2019.

[6] A. A. Silva, S. Gupta, A. M. Bazzi, and A. Ulatowski, "Wavelet-based information filtering for fault diagnosis of electric drive systems in electric ships," *ISA Transactions*, vol. 78, pp. 105–115, 2018.

Research Article

Network-Based Pharmacological Study on the Mechanism of Action of Buxue Liqi Huatan Decoction in the Treatment of Lung Cancer

Huabing Wei, Lihuang Zhou, Xiaojing Zhao ⓘ, and Feng Xie ⓘ

Department of Thoracic Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

Correspondence should be addressed to Xiaojing Zhao; zhaoxiaojing@renji.com and Feng Xie; jiefeng@renji.com

Objective. This study aimed to investigate the mechanism of action of Buxue Liqi Huatan decoction against lung cancer through network pharmacology. *Methods.* The chemical composition and targets of all the drugs in the Buxue Liqi Huatan decoction were obtained through the Database and Systematic Analysis Platform of Traditional Chinese Medicine Pharmacology, the Integrated Database of Traditional Chinese Medicine, and by screening lung cancer targets with the gene map and OMIM database. The targets were then imported into Cytoscape 3.7.2 to build a target network of active ingredients and imported into the STRING database to build a protein-protein interaction network. The BisoGenet plug-in in Cytoscape 3.7.2 was used for network topology analysis. Genetic ontology (GO) enrichment analysis and Kyoto Encyclopedia of Gene and Genomes (KEGG) enrichment analysis were performed on potential targets of the Buxue Liqi Huatan decoction for lung cancer using the R-language Bioconductor platform, and results were imported from Cytoscape 3.7.2 to obtain the KEGG network connection diagram via the Autodock molecular docking software. *Results.* A total of 238 chemical components and 694 disease targets were obtained, including 133 intersecting targets. The key targets included TP53, AKT1, and MYC, and the GO functional analysis was mainly related to oxidative and cellular oxidative stress, apoptotic signaling, and antibiotic response. The results showed that the key target with the best binding performance was TP53. *Conclusion.* The treatment of lung cancer with blood-supplementing, qi-transforming, and phlegm-transforming soups works through multiple components and targets. The active ingredients include quercetin, luteolin, naringenin, and baicalein. Among them, the core proteins of PPI protein interaction mainly include TP53, AKT1, MYC, EGRF, CCNB1, and ESR1. The enrichment analysis results show that the TNF signal pathway, PI3K-Akt signal pathway, AGE-RAGE, IL-17, etc., are the main signal pathways of Buxue Liqi Huatan decoction in treating lung cancer. This lays the foundation for further study of its mechanism.

1. Introduction

Lung cancer is one of the most common and fatal malignant tumors in China [1], and the five-year survival rate of patients with lung cancer is only 19% [2]. According to statistics, in 2018, more than 2.9 million new cases of lung cancer were reported worldwide, and the death toll from lung cancer reached more than 1.76 million. The incidence of lung cancer in men is higher than that in women, and it is higher in urban areas than in rural areas; therefore, the incidence and mortality rates have significant regional differences [3]. The surgical resection and chemotherapy are often used for the clinical treatment of lung cancer; however, the quality of life of the patient becomes low after these. Traditional Chinese Medicine (TCM) treatment based on syndrome differentiation plays an important role in the management of precancerous lesions and the prevention of tumor metastasis and multidrug resistance [4].

Network pharmacology explains the occurrence and development of diseases from the perspectives of system biology and biological network balance, recognizes the interactions between drugs, and guides the discovery of new drugs to improve or restore balance to the biological network of the body. In recent years, an increasing number of

researchers have explored the relationship between composite multielement targets and diseases from the perspective of system biology, thereby providing a new direction for modern TCM research [5]. Therefore, in this study, the pharmacology of the network was used as a breakthrough to analyze the target action of Buxue Liqi Huatan decoction in the treatment of lung cancer and predict the possible mechanisms underlying its use.

2. Materials and Methods

2.1. Screening of Target Components of Buxue Liqi Huatan Decoction.
The active chemical components of Chinese medicines (*Rhizoma Pinelliae*, *Rhizoma Arisaematis*, *Tianlong*, *Scorpio*, *Pericarpium Citri Reticulate*, *Endothelium Corneum Gigeriae Galli*, and *Radix Glycyrrhizae Preparata*) in the Buxue Liqi Huatan decoction were searched through TCM system pharmacology platforms: Traditional Chinese Medicine Pharmacology (TCMSP) and the Integrated Database of Traditional Chinese Medicine (TCMID). TCMSP was selected based on these settings: oral biologics (OB) ≥30% and drug dependence (DL) ≥0.18 [5]. TCMID was screened using SwissADME. The screening conditions were as follows: a high gastrointestinal (GI) absorption rate (pharmacokinetics) and more than two "yes" in drug-likeness. Ineffective components were removed, and the active components of Chinese medicine in the Buxue Liqi Huatan decoction were obtained. Potential proteins were obtained through the Swiss forecast of the objectives database with a screening condition of a probability ≥0.1, and the screened protein targets were converted to standardized gene names in the UniProt database.

2.2. Screening of Lung Cancer Disease-Related Targets.
Lung cancer-related target genes were obtained using "gastric cancer" as the search term in the OMIM and GeneCards databases. There were too many targets retrieved from the GeneCards database, and the number of targets was filtered based on score values. A larger score value indicated that the target had a stronger correlation with the disease. Based on the assumption that a target with a score value greater than the median was the potential target for the disease, genes with scores greater than 5 that were left behind in the GeneCards database were merged and deduplicated in the OMIM database, that is, the target genes that were related to lung cancer disease.

2.3. Acquisition of Effective Targets and Venn Diagram.
A Venn diagram was used to determine the intersection of the target of the Buxue Liqi Huatan decoction and the target of lung cancer to obtain the intersection target of the two, which was the effective target of Buxue Liqi Huatan decoction in the treatment of lung cancer.

2.4. Analysis of Active Ingredient-Effective Target Network Construction Analysis.
Drug active components and effective target genes were imported into Cytoscape 3.7.2

software for networking and visual analysis [6], and a drug active components-target network diagram was obtained. The importance of the drug active components and their action targets was evaluated by analyzing network topology parameters, such as the degree value.

2.5. Protein Network Construction.
The effective targets of Buxue Liqi Huatan decoction and lung cancer were imported into the STRING database to create a protein interaction network, and data with a reliability ≥0.900 were selected. Results were imported into Cytoscape 3.7.2 software for visual analysis in the tsv format. Protein-protein interaction (PPI) data were imported using R software to learn the connection points of core genes [6], and a histogram showing the top 30 core genes was obtained.

2.6. Enrichment Analysis of Target Function and Pathway.
The software R (https://www.r-project.org/) and its base database were used to obtain the genetic ID (entrezid) of the potential action point. Other programs such as DOS, Cluster Profiler, and others were used as well. The Pathview (bioconductor) software package was used to analyze the GO function of these potential targets, including biological process (BP), cell component (CC), and molecular function (MF) analyses. The p value cutoff was 0, and the q-value cutoff was 0.05. GO concentration analysis was divided into three categories: BP, MF, and CC. Each category was classified according to the degree of importance, and the first 10 rich items are presented in a histogram and bubble chart.

2.7. Main Active Components-Target Molecular Docking of Buxue Liqi Huatan Decoction.
The targets of Buxue Liqi Huatan decoction on lung cancer were searched in the protein data bank (PDB) database and saved in PDB format. The ligands were stored in the mol2 format as the top 2 compounds by degree after topological analysis. The potential targets of Buxue Liqi Huatan decoction on lung cancer were molecularly docked with the main compounds in Buxue Liqi Huatan decoction using AutoDockTools 1.5.6.

3. Results

3.1. Acquisition of Active Components and Related Objectives of the Buxue Liqi Huatan Decoction.
In TCMSP and TCMID, 238 active components in Buxue Liqi Huatan decoction were obtained using OB ≥30% and DL ≥0.18, a high GI absorption (pharmacokinetics), and more than two "yes" in drug-likeness as the screening conditions. The active components are shown in the supporting file (Table S1, supporting information). A total of 654 Chinese medicine targets were retrieved through TCMSP and TCMID.

3.2. Acquisition of Lung Cancer-Related Targets.
Disease genes were obtained from GeneCards and OMIM databases. Based on experience, targets with scores greater than the median were set as potential disease targets, and the related targets retrieved from the OMIM database were combined,

followed by the deletion of duplicate values. Finally, 694 lung cancer-related targets were obtained.

3.3. Venn Drawing. The Venn tool in TBtools was used to intersect the targets of Buxue Liqi Huatan decoction with those of lung cancer, and 133 intersecting targets of the two were obtained, with the result shown in Figure 1.

3.4. Construction of Network Diagram of Active Components-Effective Targets of Buxue Liqi Huatan Decoction. The active components and effective target network of Buxue Liqi Huatan decoction were constructed using Cytoscape 3.7.2, as shown in Figure 2. The topological parameters of the network of Buxue Liqi Huatan decoction for the treatment of lung cancer were calculated using the software, to assess the importance of active components and action objectives. Results showed that the active components, such as *quercetin*, *luteolin*, and *naringenin*, could act on multiple targets, and these components might be the main active components of Buxue Liqi Huatan decoction in the treatment of lung cancer.

3.5. Protein Network Construction. The Venn tool in TBtools was used to intersect the target of Buxue Liqi Huatan decoction with the target of lung cancer, as shown in Figure 1. The intersected target was uploaded to the STRING database, and the setting credibility of ≥0.9 was used to obtain a target PPI network diagram. Data were imported into Cytoscape 3.7.2 to draw a protein network diagram. Larger degree values correspond to larger nodes. The location in the network was determined based on the degree value. According to Figure 3, the targets in the network center were TP53, AKT1, and MYC, which were presumed to be important targets of Buxue Liqi Huatan decoction for the treatment of lung cancer.

3.6. Results of Enrichment Analysis of Target Function and Pathway. Using R to perform GO annotation analysis on effective targets, we selected the top 10 results from BP, CC, and MF and found that BP was primarily involved in oxidative stress, cell oxidative stress, apoptosis signal, antibiotic response, and other aspects. MF is primarily involved in ubiquitination protein ligase binding, cytokine corresponding binding, phosphorylation site binding, etc. CC is primarily associated with nuclear chromatin, membrane rafts, transcription factor complexes, and the microfiltration zone. The results are shown in Figure 4(a). After KEGG enrichment analysis, 154 pathways were identified, including the AGE-RAGE, tumor necrosis factor (TNF), P13K-Akt, and IL-17 pathways. The first 20 pathways were selected and visualized, and the results are shown in Figure 4(b).

3.7. Molecular Docking Results of the Active Components in Buxue Liqi Huatan Decoction. AutoDockTools 1.5.6 was used to conduct molecular coupling between potential targets of Buxue Liqi Huatan decoction on lung cancer and
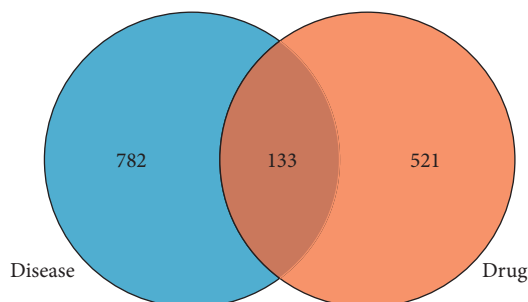
FIGURE 1: Venn diagram of the intersection target of Buxue Liqi Huatan decoction and lung cancer.

the main compounds after topology analysis and calculation. The more stable the structure of receptors and receptor combinations, the higher the possibility of occurrence. The top 2 targets were selected according to the degree value, namely, AKT1 and TP53 tumor protein. AutoDock Vina was used for molecular docking of the components, and a binding energy <−4.25 kcal/mol indicated that there was a certain binding activity between the small ligand molecule and the receptor protein. Binding energy <−5.0 kcal/mol indicated that there was good binding activity between them. Binding energy <−7.0 kcal/mol indicated that the ligand had a strong binding activity with the receptor [6]. The binding energies of quercetin core targets were −6.7 kcal/mol and −7.7 kcal/mol, indicating that the drug had a strong binding activity with the target paper. The specific docking results are shown in Figure 5.

4. Discussion

In this study, 238 active components were screened using network pharmacology, corresponding to 654 active component action targets, which were intersected with 915 gene targets. A total of 133 common targets were obtained, among which the active components included quercetin, luteolin, naringenin, and baicalein. The screened compounds were consistent with the results of existing studies, indicating that the screening results had a good reference value.

Quercetin has been proven to reduce oxidative stress, interfere with the renin-angiotensin system, and downregulate the reactive oxygen downstream signaling pathway to exert antioxidant, antitumor, anti-inflammatory, antibacterial, and cardiovascular protection as well as other pharmacological effects [7]. Studies have shown that luteolin has multitarget, multipathway, and multilink antitumor activities. Its mechanism of action includes inhibition of tumor cell proliferation, induction of tumor cell apoptosis, and inhibition of tumor metastasis [8]. A helmet can enhance the killing activity of natural killer cells on cancer cells, prevent the growth and migration of cancer cells, lead to the death of cancer cells, and hinder the glucose metabolism of cancer cells [9]. Baicalein can inhibit cell cycle progression and thus inhibit the proliferation of tumor cells by regulating the levels of different types of cyclins and cell cycle-dependent kinases [10]. In this study, we found that quercetin, luteolin, and baicalein could be connected to more protein
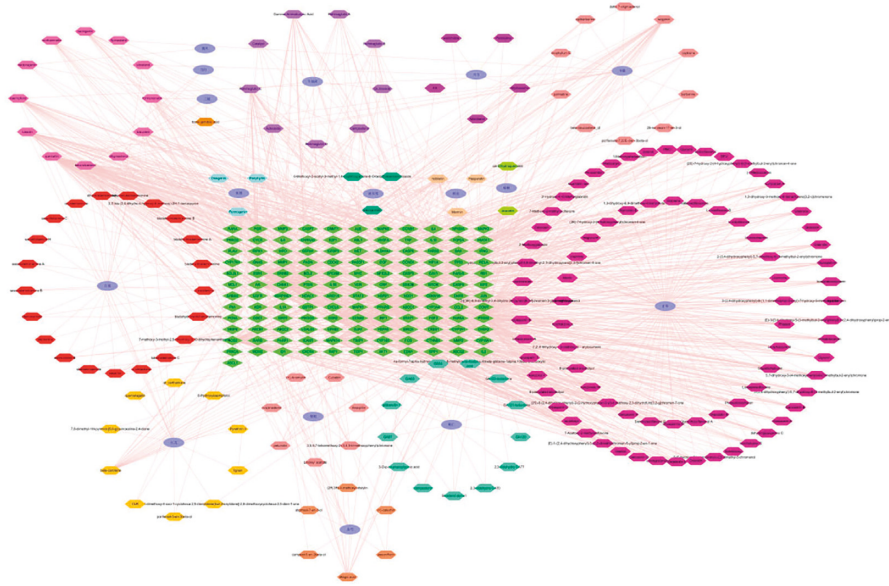
FIGURE 2: Network diagram of active components and effective targets of Buxue Liqi Huatan decoction.
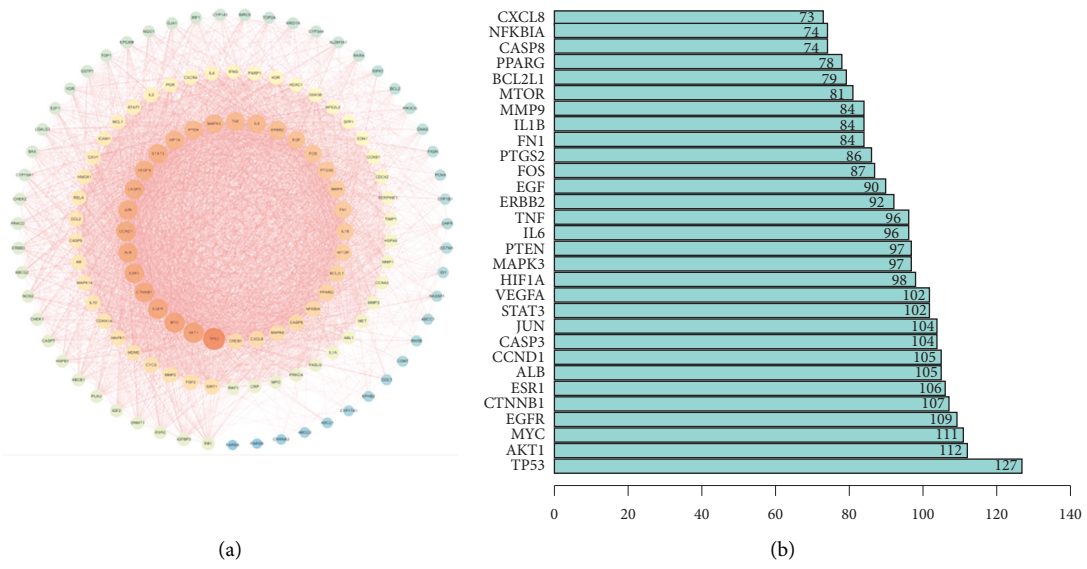
(a)

(b)
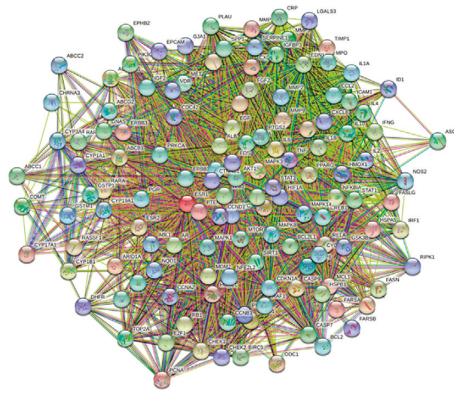
(c)

FIGURE 3: PPI network diagram of intersection targets. (a) PPI network diagram, (b) core gene map, and (c) protein mutual mapping.
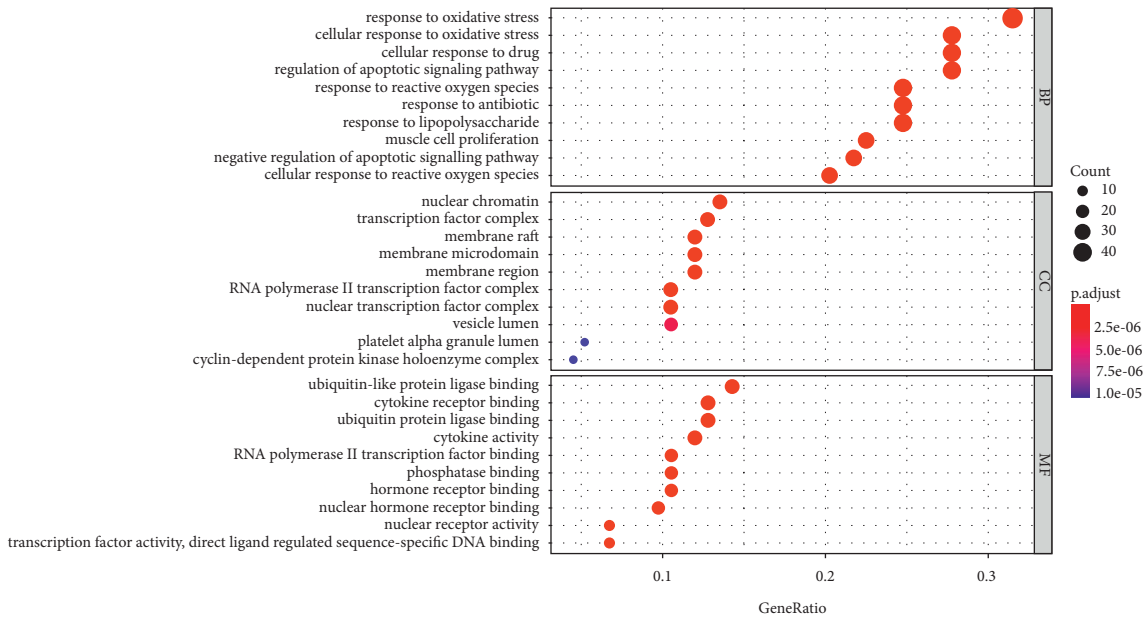
(a)
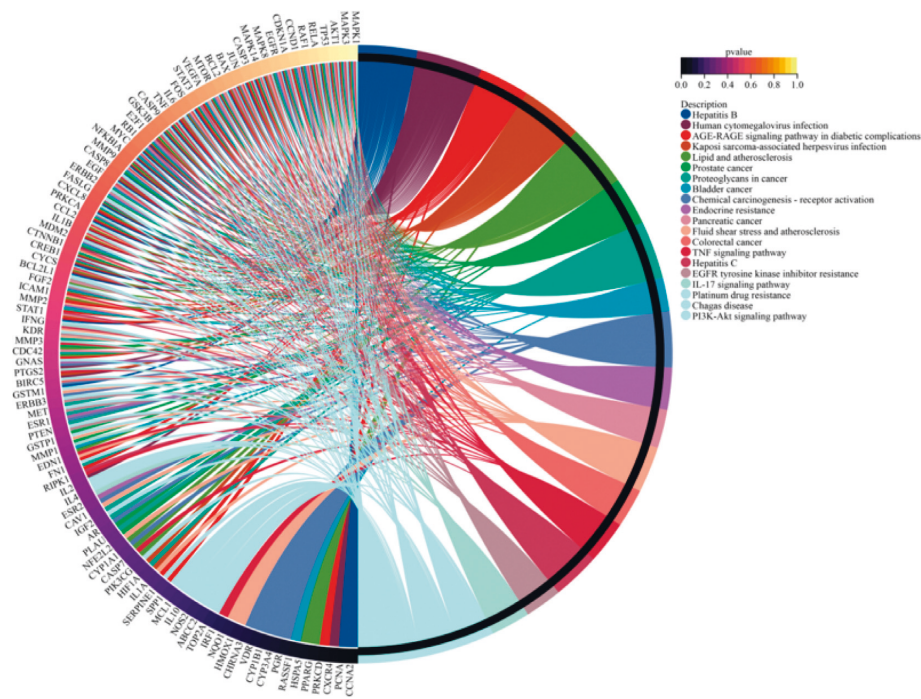
(b)

FIGURE 4: Analysis of enrichment of lung cancer treated with Buxue Liqi Huatan decoction. (a) GO enrichment analysis and (b) KEGG enrichment analysis circle.

targets and had good docking activity with the core targets, suggesting that they have potential anticancer activity. Among them, the core proteins for PPI were TP53, AKT1, MYC, eGFR, CCNB1, and ESR1. TP53 is a tumor suppressor gene, which is responsible for regulating the proliferation of tumor cells. The TP53 gene is an important anticancer gene, and its wild type can cause cancer cell apoptosis, thus preventing cancer, and it also has the function of helping cell genes repair defects [11, 12]. As an oncogene, AKT1 expression is upregulated in lung cancer. A study has revealed

that the expression of AKT1 protein is related to the differentiation degree of lung cancer, lymph node metastasis, and TNM stage. A higher expression indicates that the lung cancer has a lower differentiation degree, a higher malignancy degree, and a worse prognosis, indicating that it is related to the procancer effect in the development of lung cancer [13]. Luteolin has a good binding ability to AKT1, can inhibit the proliferation, invasion, and migration of non-small cell lung cancer (NSCLC) A549 cells, and can induce apoptosis [14]. By regulating the expression of troponin and

(a)

(b)
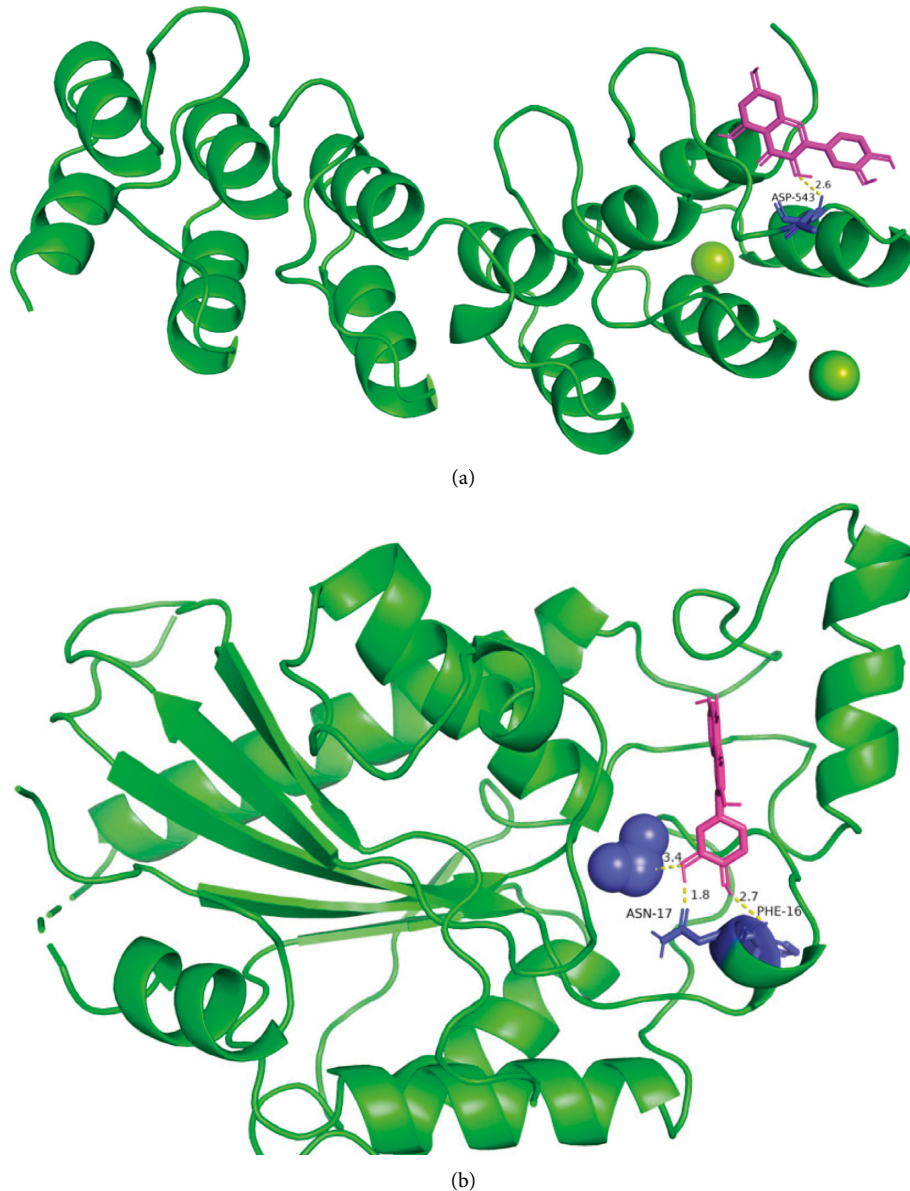
Figure 5: Molecular docking mode diagram. (a) Docking diagram of quercetin and AKT1 and (b) docking diagram of quercetin and TP53.

shrinking the cytoskeleton to reverse epithelial-mesenchymal transition (EMT), baicalein can inhibit the growth of lung cancer cells through the targeted Src/Id1 pathway. In addition, it can enhance the toxicity of cisplatin and enhance the sensitivity to chemotherapy, which may be related to EMT inhibition by blocking the intracellular Akt/NF-κB signaling pathway [15]. The Myc gene was found in Burkitt lymphoma for the first time and can be activated by chromosome translocation [16]. Overexpression of myc is found in various tumors, such as lung cancer, stomach cancer, breast cancer, colon cancer, cervical cancer, some neuroblastomas, granulocytic leukemia, retinoblastoma, osteosarcoma, chondrosarcoma, chordoma, liposarcoma, rhabdomyosarcoma, Hodgkin's disease, and head tumors. All of them have amplification or overexpression of the myc gene. At least three Myc genes have been found, including C-myc, N-myc, and L-myc [17, 18].

The KEGG pathway enrichment analysis showed that TNF, PI3K-Akt, AGE-RAGE, and IL-17 signaling pathways were the main signaling pathways of Buxue Liqi Huatan decoction in the treatment of lung cancer. Among them, the PI3K-Akt signaling pathway plays an important role in tumor diseases [19, 20]. Overactivation of the PI3K/Akt signaling pathway leads to a decrease in the expression of tumor suppressor protein p53, which promotes the synthesis of protein, the proliferation of tumor cells, and inhibition of cell apoptosis [19]. Studies have found that the AGE/RAGE signaling pathway plays a role in the activation of EMT in cancer cells, and it can affect the dryness of cancer cells by regulating EMT. In addition, the PI3K/Akt signaling pathway can be used as the pathway between AGE/RAGE and EMT [21]. IL-17-mediated tumor angiogenesis is involved in the activation of the Stat3/GIV signaling pathway and subsequent upregulation of vascular endothelial growth

factor production in NSCLC cells. This may be the mechanism underlying the correlation between IL-17, poor prognosis, and angiogenesis. Therapies targeting IL-17 and GIV can be developed as potential treatments to inhibit lung cancer [22, 23].

TNF is a central proinflammatory cytokine that is involved in a variety of inflammatory states, including autoimmunity. In 1975, Carswell et al. found that after BCG-vaccinated mice were injected with bacterial lipopolysaccharide, a substance appeared in the serum that could cause hemorrhagic necrosis of various tumors, and it was named tumor necrosis factor (TNF) [24]. TNF can kill some tumor cells in vivo and in vitro or inhibit proliferation. The sensitivity of tumor cell lines to TNF-α is very different, and TNF-α can even stimulate a few tumor cells [25, 26]. Treating tumor cells (such as mouse fibroblast L929) with actinomycin D, mitomycin C, and cycloheximide can obviously increase the killing activity of TNF-α on tumor cells. The response of tumors to TNF-α in vivo is also very different, which is not parallel to the sensitivity of tumor cell lines to TNF-α in vitro. There may be sensitive and resistant strains such as L929-S and L929-R in the same cell line. In addition, the expression of endogenous TNF in target cells may make cells resist the cytotoxic effect of exogenous TNF, so the sensitivity of cells to exogenous TNF can be changed by inducing or inhibiting the expression of endogenous TNF [27, 28]. Macrophage-binding TNF may be involved in killing target cells. TNF can affect various effector factors of immune monitoring and cause enhanced activity of natural killer cells and cytotoxic T lymphocytes [29]. TNF can also directly affect tumors and drive cells to respond to apoptosis, necrosis, survival, inflammation, or growth promotion. TNFR1 and TNFR2 are associated with various intracellular signaling pathways. TNFR1 can induce the expression of genes related to inflammation, survival, and proliferation by activating kinase signals and finally activating the nuclear factor NF-kB [30–32].

In summary, based on network pharmacology, we preliminarily explored the possible mechanisms underlying Buxue Liqi Huatan decoction in the treatment of lung cancer. The predicted action targets in this study are consistent with those of existing studies on Buxue Liqi Huatan decoction in the treatment of lung cancer, therefore reflecting the accuracy of target prediction. In addition, this study revealed that Buxue Liqi Huatan decoction treated lung cancer based on its multicomponent, multitarget, and multipathway activities and further provided the basis for experimental verification and clinical research. This study had some limitations. The targets obtained from different databases were not identical, so there might be discrepancies. Different screening criteria were adopted for data processing. Therefore, further experimental studies are required to verify and supplement the results of this study.

Data Availability

The datasets used and analyzed in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Authors' Contributions

Xiaojing Zhao and Feng Xie designed and directed the study, Huabing Wei and Lihuang Zhou collected all the data and wrote the paper. All authors have read and approved the final manuscript. Huabing Wei and Lihuang Zhou have contributed equally to this work.

Acknowledgments

Supplementary Materials

In the TCMSP and TCMID databases, 238 active components in Buxue Liqi Huatan decoction were obtained using OB ≥30% and DL ≥0.18, the active components are shown in Table S1. A total of 654 Chinese medicine targets were retrieved through TCMSP and TCMID databases. (*Supplementary Materials*)

References

[1] F. Jacques, C. Murielle, S. Isabelle et al., "Cancer statistics for the year 2020: an overview," *International Journal of Cancer*, vol. 149, no. 4, pp. 12–14, 2021.

[2] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022," *CA: A Cancer Journal for Clinicians*, vol. 72, no. 1, pp. 7–33, 2022.

[3] W. Chen, R. Zheng, P. D. Baade et al., "Cancer statistics in China, 2015," *CA: A Cancer Journal for Clinicians*, vol. 66, no. 2, pp. 115–132, 2016.

[4] T. Ya. Kucherova, V. V. Velikaya, O. V. Gribova et al., "Physical therapy methods in the treatment and rehabilitation of cancer patients," *AIP Conference Proceedings*, vol. 1760, no. 1, p. 198, 2016.

[5] X. Wang, Zi-Yi Wang, J. H. Zheng, and S. Li, "TCM network pharmacology: a new trend towards combining computational, experimental and clinical approaches," *Chinese Journal of Natural Medicines*, vol. 19, no. 1, p. 11, 2021.

[6] C. Yuan, M.-H. Wang, F. Wang et al., "Network pharmacology and molecular docking reveal the mechanism of Scopoletin against non-small cell lung cancer," *Life Sciences*, vol. 270, 2021.

[7] D. Martina, V. Miroslav, S. Ernest, and J. Soňa, "Antioxidant action and cytotoxicity on HeLa and NIH-3T3 cells of new quercetin derivatives," *Interdisciplinary Toxicology*, vol. 6, no. 4, p. 11, 2013.

[8] Y. Pu, T. Zhang, J. Wang et al., "Luteolin exerts an anticancer effect on gastric cancer cells through multiple signaling pathways and regulating miRNAs," *Journal of Cancer*, vol. 9, no. 20, pp. 3669–3675, 2018.

[9] Y. Gao, S. A. Snyder, J. N. Smith, and Y. C. Chen, "Anticancer properties of baicalein: a review," *Medicinal Chemistry Research*, vol. 25, no. 8, pp. 1515–1523, 2016.

[10] M. Jiasheng and L. Tianrun, "The traditional Chinese medicine baicalein potently inhibits gastric cancer cells," *Journal of Cancer*, vol. 7, no. 4, 2016.

[11] B. J. Aubrey, A. Strasser, and G. L. Kelly, "Tumor-suppressor functions of the TP53 pathway," *Cold Spring Harbor Perspectives in Medicine*, vol. 6, no. 5, Article ID a026062, 2016.

[12] W. Fu, Z. J. Zhuo, W. Jia et al., "Association between TP53 gene Arg72Pro polymorphism and Wilms’ tumor risk in a Chinese population," *OncoTargets and Therapy*, vol. 10, pp. 1149–1154, 2017.

[13] R. Mohd and M. Maged, "Virtual screening of 1, 4-Naphthoquinone derivatives for inhibition of a key cancer signaling protein, AKT1 kinase," *Anticancer Research*, vol. 39, no. 7, 2019.

[14] X. Yao, W. Jiang, D. Yu, and Z. Yan, "Luteolin inhibits proliferation and induces apoptosis of human melanoma cells *in vivo* and *in vitro* by suppressing MMP-2 and MMP-9 through the PI3K/AKT pathway," *Food & Function*, vol. 10, no. 2, pp. 703–712, 2019.

[15] X. Yu, Y. Yang, Y. Li et al., "Baicalein inhibits cervical cancer progression via downregulating long noncoding RNA BDLNR and its downstream PI3 K/Akt pathway," *The International Journal of Biochemistry & Cell Biology*, vol. 94, pp. 107–118, 2018.

[16] R. T. Maguire, T. S. Robins, S. S. Thorgeirsson, and C. A. Heilman, "Expression of cellular myc and mos genes in undifferentiated B cell lymphomas of Burkitt and non-Burkitt types," *Proceedings of the National Academy of Sciences*, vol. 80, no. 7, pp. 1947–1950, 1983.

[17] F. R. Dejure and M. Eilers, "MYC and tumor metabolism: chicken and egg," *The EMBO Journal*, vol. 36, no. 23, pp. 3409–3420, 2017.

[18] G. M. Ramsay, G. Moscovici, C. Moscovici, and J. M. Bishop, "Neoplastic transformation and tumorigenesis by the human protooncogene MYC," *Proceedings of the National Academy of Sciences*, vol. 87, no. 6, pp. 2102–2106, 1990.

[19] L. Yu, W. Jessica, and P. Liu, "Attacking the PI3K/Akt/mTOR signaling pathway for targeted therapeutic treatment in human cancer," *Seminars in Cancer Biology*, vol. S1044-579X, no. 21, 2021.

[20] B. M. Slomovitz and R. L. Coleman, "The PI3K/AKT/mTOR pathway as a therapeutic target in endometrial cancer," *Clinical Cancer Research*, vol. 18, no. 21, pp. 5856–5864, 2012.

[21] S.-Y. Ko, H.-An Ko, T. M. Shieh et al., "Cell migration is regulated by AGE-RAGE interaction in human oral cancer cells in vitro," *PLoS One*, vol. 9, no. 10, Article ID e110542, 2014.

[22] D. He, H. Li, N. Yusuf et al., "IL-17 promotes tumor development through the induction of tumor promoting microenvironments at tumor sites and myeloid-Derived suppressor cells," *The Journal of Immunology*, vol. 184, no. 5, pp. 2281–2288, 2010.

[23] B. Pan, J. Shen, J. Cao et al., "Interleukin-17 promotes angiogenesis by stimulating VEGF production of cancer cells via the STAT3/GIV signaling pathway in non-small-cell lung cancer," *Scientific Reports*, vol. 5, no. 1, p. 16053, 2015.

[24] E. A. Carswell, L. J. Old, R. L. Kassel, S. Green, N. Fiore, and B. Williamson, "An endotoxin-induced serum factor that causes necrosis of tumors," *Proceedings of the National Academy of Sciences*, vol. 72, no. 9, pp. 3666–3670, 1975.

[25] Y. Wu and B. P. Zhou, "TNF-α/NF-κB/Snail pathway in cancer cell migration and invasion," *British Journal of Cancer*, vol. 102, no. 4, pp. 639–644, 2010.

[26] T. Schioppa, R. Moore, R. G. Thompson et al., "B regulatory cells and the tumor-promoting actions of TNF-α during squamous carcinogenesis," *Proceedings of the National Academy of Sciences*, vol. 108, no. 26, pp. 10662–10667, 2011.

[27] M. Grell, G. Zimmermann, E. Gottfried et al., "Induction of cell death by tumour necrosis factor (TNF) receptor 2,CD40 and CD30: a role for TNF-R1 activation by endogenous membrane-anchored TNF," *The EMBO Journal*, vol. 18, no. 11, pp. 3034–3043, 1999.

[28] L. Yang, Y. Wang, S. Shi et al., "The TNF-α-induced expression of miR-130b protects cervical cancer cells from the cytotoxicity of TNF-α," *FEBS Open Bio*, vol. 8, no. 4, pp. 614–627, 2018.

[29] L. Zhu, S. Hu, Q. Chen et al., "Macrophage contributes to radiation-induced anti-tumor abscopal effect on transplanted breast cancer by HMGB1/TNF-α signaling factors," *International Journal of Biological Sciences*, vol. 17, no. 4, pp. 926–941, 2021.

[30] P. Xia, R. Zhang, and G. Ge, "C/EBPβ mediates TNF-α-induced cancer cell migration by inducing MMP expression dependent on p38 MAPK," *Journal of Cellular Biochemistry*, vol. 116, no. 12, pp. 2766–2777, 2015.

[31] W. Cai, Z. J. Kerner, H. Hong, and J. Sun, "Targeted cancer therapy with tumor necrosis factor-alpha," *Biochemistry Insights*, vol. 2008, no. 1, pp. 15–21, 2008.

[32] F. Flego, R. B. Radojcić Badovinac, B.-K. Bulat-Kardum et al., "Primary lung cancer and TNF-alpha gene polymorphisms: a case-control study in a Croatian population," *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, vol. 15, no. 7, p. CR361, 2009.

Research Article

Epstein–Barr Virus (EBV) and Multiple Sclerosis Disease: A Biomedical Diagnosis

Asma Alanazi [ID] [1,2]

[1]*Department of Basic Medical Sciences, College of Medicine, King Saud Bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia*
[2]*King Abdullah International Medical Research Centre (KAIMRC), Riyadh, Saudi Arabia*

Correspondence should be addressed to Asma Alanazi; anazia@ksau-hs.edu.sa

Multiple sclerosis (MS) is a degenerative disease that affects 2.8 million people worldwide. It is a central nervous system disease (CNS), in which the myelin sheath covering the brain and spinal cord neurons is attacked. If the myelin sheath is damaged, a person can suffer permanent damage to the nerves. There are a number of factors that can increase a person's risk of developing MS, such as obesity, smoking, vitamin D deficiency, certain tissue types (HLADRB1 ∗ 15 : 01) and infection with the Epstein–Barr virus (EBV). The latter virus can cause infectious mononucleosis, which can, in turn, result in lifelong infection in the host. To establish the relationship between MS and EBV, the author conducted a study on 1176 MS patients admitted to Saudi Arabia King Abdulaziz City centers. The researcher determined that MS occurred twice as much in females as it did in males, and also that EBV was much more widespread in MS female patients than MS male patients (27 : 1). Age was not a factor in the occurrence of EBV. There were limitations on data completeness and availability. Other trials using larger cohorts of patients are needed.

1. Introduction

Multiple sclerosis (MS) is a chronic inflammatory disease impacting the central nervous system (CNS) and one of the leading causes of neurological disability in adults. It is caused by the immune cells that attack CNS-self cells as if they were foreign bodies, which causes demyelination and neuro-degeneration in the CNS. The damage leads to a gradual deterioration of sensory, motor, and cognitive functions [1, 2]. Although there are various forms of MS, the disease typically takes the form of relapsing-remitting (RR) cycles of disease activity. In rare cases, primary-progressive (PP) disease occurs. It is common for RRMS to progress over time and become a secondary-progressive (SP) disease [3]. Not everyone is susceptible to MS; the host must be genetically susceptible in the HLA-II locus after interaction with environmental factors [4]. MS arises from interactions between an individual's genetic susceptibility and environmental factors. One of the most MS-associated susceptibility alleles is HLA DRB1 ∗ 1501 [5]. Moreover, there are several

environmental risk factors that can increase a person's risk of developing MS, such as exposure to sunlight/high vitamin D levels, smoking, and infectious agents [1]. One of the most common infectious agents that causes the condition is the Epstein–Barr virus (EBV), a human DNA herpes virus that has a strong connection with MS [6, 7].

Research shows that EBV is one of the world's most well-developed human viruses. It is usually spread through saliva. In most infected people, EBV creates an asymptomatic, lifelong infection which is carried as an infection of B lymphocytes for the duration of one's life. The virus, however, is controlled by the host's immune system. It can–very rarely though–cause infectious mononucleosis and human tumor malignancies, such as lymphomas and nasopharyngeal carcinoma [8]. The rare occasions detailed earlier are due to a virus-host imbalance, one that brings about the pathogenic potential of EBV [9, 10].

Two studies conducted in Gulf Cooperation Council countries, as well as one study in Kuwait and a further study, have examined the relationship between EBV and MS. used
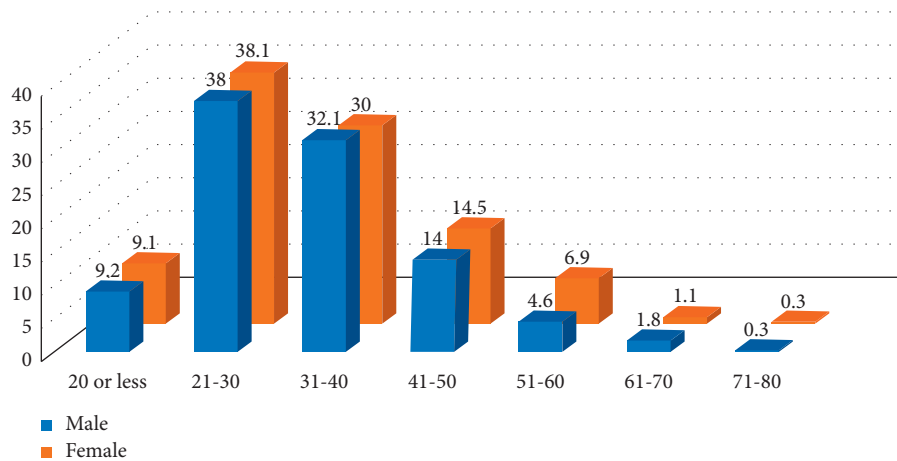
Figure 1: Distribution of MS by age group in a sample of 1176 patients.

data obtained from US military recruits over a period of 20 years between 1993 and 2013 [11]. The studies had conflicting reports. One study in Kuwait was epidemiological and did not establish an association between EBV and MS, while the other one in Kuwait was a serological study that showed a strong link between EBV and MS [12]. What's more, [11] found the risk of developing MS is largely increased by contracting EBV infection. Moreover, it was found that this virus often precedes the onset of MS, which further supports the belief that it may play a role in MS pathogenesis. The study involved over 10 million young adults working for the US army. Risks of MS increased 32-fold post-EBV infection, but not other viruses of similar infection method, such as the cytomegalovirus. The findings showed that, after EBV seroconversion occurred, there was an increase in the serum levels of the neurofilament light chain, which is a critical indicator of neuroaxonal degeneration. Other than the presented studies, a number of large-scale epidemiological and serological studies in Western countries showed that EBV is strongly associated with MS [13], evidence being virtually 100% EBV seroprevalence in adult MS, and higher anti-EBV antibody titers, specifically for EBNA1 IgG in MS patients [7, 14], and an increased risk of developing MS after infectious mononucleosis [15–17]. Infectious mononucleosis, a feature of EBV pathology, more than doubles the risk of MS.

In EBV-infected individuals, the immune reactivity toward the virus is higher in people with MS, indicating inadequate control of the virus [15]. Meanwhile, evidence from neuropathological and immunological studies indicates that continual EBV infection in the central nervous system can generate an immunopathological CD8 T-cell response in order to eliminate the virus. However, this ultimately results in bystander CNS injury [18]. We searched for EBV-specific CD8 T cells in MS brain tissue. The latter researchers examined the postmortem brain samples of 12 donors who had progressive MS and had been diagnosed with the HLA class 1 genotype [19]. Their work concluded that the expression of CD8 cytotoxic T cells was detected toward the expression of proteins in the latent and lytic phases of the EBV life cycle. In addition, they determined that these immune cells were
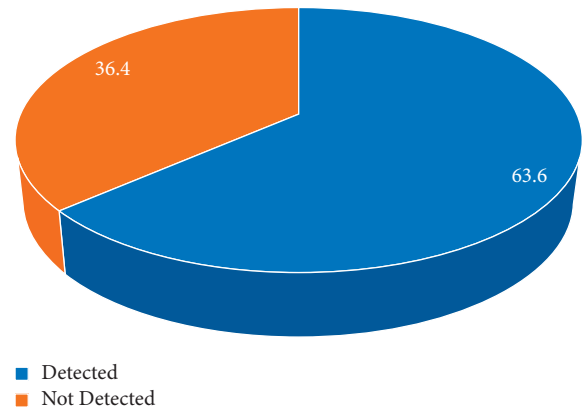
Figure 2: Percent of EBV detected in a sample of 44 MS patients.

detected in the white matter lesions and/or meninges of 11–12 MS donors. The T cells were also found to be adhered to EBV-infected cells. The median value of CD8 T cells recognizing individual EBV epitotes fell in the range of 0.5 to 2.5% of CNS-infiltrating CD8 T cells. CNS-infiltrating EBV-specific CD8 T cells showed a cytotoxic phenotype as they were CD107a positive. Additionally, the relationship between EBV-related immune responses and CNS-related inflammation that causes MS pathology is supported by the presence of local EBV dysregulation and selective enrichment of EBV-specific CD8 T cells in the brains of MS sufferers. the prevalence of MS has increased in Saudi Arabia, and there is no study in Saudi Arabia evaluating the correlation between EBV and MS. Therefore, the aim of the study is to assess the association between EBV and MS patients in King Abdulaziz medical centers from 2015 to 2022.

2. Methods

A descriptive cross-sectional study was carried out among 1176 MS patients whose electronic medical records were kept in King Abdulaziz City Centers for the admittance years of 2015 to 2022. The data were obtained by the BEST Care system. Data were coded for entry and analysis using SPSS statistical software (version 22.0). The research earned the

Table 1: Comparison between gender age of the sample of MS patients ($n = 1176$).

Variables	Number	Percent	Number	Percent
Gender				
	Male		Female	
	392	33.3	784	66.7
Age				
20 or less	36	9.2	71	9.1
21–30	149	38.0	299	38.1
31–40	126	32.1	235	30.0
41–50	55	14.0	11454	14.5
51–60	18	4.6	9	6.9
61–70	7	1.8	2	1.1
71–80	1	0.3		0.3
Mean sd	32.95 ± 10.762		33.12 ± 10.791	
Independent T test		$F = 0.681$	d$f = 1174$	$P = 0.432$

Table 2: Comparison between Male and Female in the MS-EBV cohort ($n = 28$).

	Male		Female		Chi square value	P
	Number	Percent	Number	Percent		
Detected	1	5.9	27	100	39.933	0.001
Not detected	16	94.1	0	0		

Table 3: Type of test used for detected cases ($n = 28$).

Test type	Frequency	Percent
EBNA antigen IgG	6	21.4
EBV-IgG	8	28.6
Epstein–Barr virus IgM	10	35.7
Epstein–Barr virus nuclear antigen IgG antibody	3	10.7
Epstein–Barr virus polymerase chain reaction (PCR) (quantitative)	1	3.6
Total	28	100.0

approval of the Institutional Review Board (IRB) from King Abdullah International Medical Research (KAIMRC). IRB[#] NRC22R/102/02. The demographic characteristics included frequency, percentage, mean, and standard deviation. For nonparametric risk factors, such as gender and symptoms, chi-square test was utilized. A one sample t-test was used to test the significance of the mean of interval and ratio variables, such as age as shown in Figures 1 and 2.

3. Results

Multiple sclerosis was as twice in the female population as in the male cohort in the sample. There was no statistically significant difference between males and females regarding the distribution of the disease by age group ($P = 0.432$). The mean age of female was 32.95 ± 10.762 and the mean age of males was 33.12 ± 10.791 as shown in Table 1.

In a subsample of 44 patients, the percentage of detected Epstein–Barr virus was 63 percent.

There was a statistically significant difference between males and females regarding the number of detected cases of EBV. The virus was detected in all women (100%) while it was only detected in 5.9% of males as shown in Table 2.

Epstein–Barr virus IgM was detected in more than one third of the subsample (35.7%). EBV-IgG was detected in

28.6% while Epstein–Barr virus PCR (Quantitative) was only detected in 3.6% of the sample. The average for EDSS score was 4.02 for the subgroups. The EDSS scale ranges from 0 to 10 in 0.5 unit increments that represent higher levels of disability. The EDSS scale varies between 0 and 10 in 0.5 unit increments, each of which denotes a higher level of disability as shown in Table 3.

4. Discussion

The author conducted this study by using electronic medical records in King Abdulaziz City for the years 2015–2022 on the number of MS patients and the cases of EBV detected. There were 1176 MS patients with females comprising 66.7% of the sample ($P = 0.432$), twice as much as the male cohort. Genetic factors are most likely to blame for the female preponderance of MS. The mean age of the female cohort (32.95 ± 10.762) overlapped with that of the male cohort (33.12 ± 10.791). In a sample of 44 MS patients who were on disease-modifying therapies (DMTs), EBV was detected in 28 of them. Only 1 male had EBV while the rest of the sample consisted only of females. In these 28 cases, IgM was detected in 35.7% of the samples while EBV PCR was only detected in 3.6% of the samples. EBV is known for its latent course of infection. About 95% of the population gets this

virus during childhood [20]. Seroconversion from negative to positive for EBV antibodies increases with age. The incidence peaks early in childhood and again around puberty, especially for females. This is coincident with the female predominance in MS [21,22]. However, it is possible for the virus to reactivate in some instances. This is not always accompanied by symptoms. However, symptoms are more likely to develop in individuals with compromised immune systems (i.e., those with MS).

The primary limitation of the current study is that the study population only included patients admitted to NGHA, Riyadh, Saudi Arabia.

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 are no conflicts of interest.

Acknowledgments

References

[1] A. Compston and A. Coles, "Multiple sclerosis," *The Lancet*, vol. 372, no. 9648, pp. 1502–1517, 2008.

[2] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajjej, and A. A. Shihab, "Applying dynamic systems to social media by using controlling stability," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4569879, 7 pages, 2022.

[3] R. Dobson and G. Giovannoni, "Multiple sclerosis – a review," *European Journal of Neurology*, vol. 26, no. 1, pp. 27–40, 2019.

[4] T. Olsson, L. F. Barcellos, and L. Alfredsson, "Interactions between genetic, lifestyle, and environmental risk factors for multiple sclerosis," *Nature Reviews Neurology*, vol. 13, no. 1, pp. 25–36, 2017.

[5] S. E. Baranzini and J. R. Oksenberg, "The genetics of multiple sclerosis: from 0 to 200 in 50 years," *Trends in Genetics*, vol. 33, no. 12, pp. 960–970, 2017.

[6] A. Ascherio, "Environmental factors in multiple sclerosis," *Expert Review of Neurotherapeutics*, vol. 13, pp. 3–9, 2013.

[7] R. M. Lucas, A. M. Hughes, M. L. J. Lay et al., "Epstein-Barr virus and multiple sclerosis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 82, no. 10, pp. 1142–1148, 2011.

[8] D. A. Lawson-Thorley, "Epstein-Barr virus: exploring the immune system," *Nature Reviews Immunology*, vol. 1, pp. 75–82, 2001.

[9] G. S. Taylor, H. M. Long, J. M. Brooks, A. B. Rickinson, and A. D. Hislop, "The immunology of Epstein-Barr virus-induced disease," *Annual Review of Immunology*, vol. 33, no. 1, pp. 787–821, 2015.

[10] M. B. Alazzam, A. T. Al-Radaideh, R. A. Alhamarnah, F. Alassery, F. Hajjej, and A. Halasa, "A survey research on the willingness of gynecologists to employ mobile health applications," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1220374, 7 pages, 2021.

[11] K. Bjornevik, M. Cortese, B. C. Healy et al., "Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis," *Science*, vol. 375, no. 6578, pp. 296–301, 2022.

[12] R. Al-Temaimi, R. Alroughani, S. Jacob, and F. Al-Mulla, "Gender influence in EBV antibody response in multiple sclerosis patients from Kuwait," *Journal of Neuroimmunology*, vol. 285, pp. 57–61, 2015.

[13] E. L. Thacker, F. Mirzaei, and A. Ascherio, "Infectious mononucleosis and risk for multiple sclerosis: a meta-analysis," *Annals of Neurology*, vol. 59, no. 3, pp. 499–503, 2006.

[14] M. B. Alazzam, H. Mansour, F. Alassery, and A. Almulihi, "Machine learning implementation of a diabetic patient monitoring system using interactive E-app," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5759184, 7 pages, 2021.

[15] A. E. Handel, A. J. Williamson, G. Disanto, L. Handunnetthi, G. Giovannoni, and S. V. Ramagopalan, "An updated meta-analysis of risk of multiple sclerosis following infectious mononucleosis," *PLoS One*, vol. 5, no. 9, Article ID e12496, 2010.

[16] M. B. Alazzam, A. T. Al-Radaideh, N. Binsaif, A. S. AlGhamdi, and M. A. Rahman, "Advanced deep learning human herpes virus 6 (HHV-6) molecular detection in understanding human infertility," *Computational Intelligence and Neuroscience*, vol. 20225 pages, 2022.

[17] M. M. Langille, A. Rutatangwa, and C. Francisco, "Pediatric multiple sclerosis: a review," *Advances in Pediatrics*, vol. 66, pp. 209–229, 2019.

[18] L. I. Levin, K. L. Munger, E. J. O'Reilly, K. I. Falk, and A. Ascherio, "Primary infection with the Epstein-Barr virus and risk of multiple sclerosis," *Annals of Neurology*, vol. 67, no. 6, pp. 824–830, 2010.

[19] B. Serafini, B. Rosicarelli, D. Franciotta et al., "Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain," *Journal of Experimental Medicine*, vol. 204, no. 12, pp. 2899–2912, 2007.

[20] D. A. Thorley-Lawson, "EBV persistence-introducing the virus," *Current Topics in Microbiology and Immunology*, vol. 390, no. Pt 1, pp. 151–209, 2015.

[21] S. K. Dunmire, K. A. Hogquist, and H. H. Balfour, "Infectious mononucleosis," *Current Topics in Microbiology and Immunology*, vol. 390, no. Pt 1, pp. 211–240, 2015.

[22] M. Filippi, A. Bar-Or, F. Piehl et al., "Multiple sclerosis," *Nature Reviews Disease Primers*, vol. 4, no. 1, p. 43, 2018.

Hindawi

Research Article

Machine Learning-Based Multimodel Computing for Medical Imaging for Classification and Detection of Alzheimer Disease

Fatemah H. Alghamedy [ID],[1] **Muhammad Shafiq** [ID],[2] **Lijuan Liu** [ID],[2] **Affan Yasin** [ID],[3] **Rehan Ali Khan** [ID],[4] **and Hussien Sobahi Mohammed** [ID][5]

[1]*Applied College, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia*
[2]*School of Artificial Intelligence, Neijiang Normal University, Neijiang, Sichuan, China*
[3]*School of Software, Tsinghua University, Beijing 100084, China*
[4]*Department of Electrical Engineering, University of Science & Technology, Bannu (28100), Pakistan*
[5]*University of Gezira, Wad Medani, Sudan*

Correspondence should be addressed to Hussien Sobahi Mohammed; hussiensobahi@uofg.edu.sd

Alzheimer is a disease that causes the brain to deteriorate over time. It starts off mild, but over the course of time, it becomes increasingly more severe. Alzheimer's disease causes damage to brain cells as well as the death of those cells. Memory in humans is especially susceptible to this. Memory loss is the first indication of Alzheimer's disease, but as the disease progresses and more brain cells die, additional symptoms arise. Medical image processing entails developing a visual portrayal of the inside of a body using a range of imaging technologies in order to discover and cure problems. This paper presents machine learning-based multimodel computing for medical imaging for classification and detection of Alzheimer disease. Images are acquired first. MRI images contain noise and contrast problem. Images are preprocessed using CLAHE algorithm. It improves image quality. CLAHE is better to other methods in its capacity to enhance the look of mammography in minute places. A white background makes the lesions more obvious to the naked eye. In spite of the fact that this method makes it simpler to differentiate between signal and noise, the images still include a significant amount of graininess. Images are segmented using the k-means algorithm. This results in the segmentation of images and identification of region of interest. Useful features are extracted using PCA algorithm. Finally, images are classified using machine learning algorithms.

1. Introduction

A single image is capable of communicating more than just one single word. When making a decision, the information that is represented visually is always given the most importance, regardless of the other information that may be accessible. The importance of image processing and the applications it has fostered over the course of the last several decades has skyrocketed in a number of different academic subfields. The proliferation of different imaging techniques has been a driving force behind the expansion of image processing as a field. The field of digital image processing is branching out into a variety of new subfields of research, one of which being medical image processing. It is a subfield of radiology in which information collected from a patient's medical imaging is analyzed in order to establish whether or not the individual in question is suffering from a disease. Inside of a person's body is where the vast majority of diseases manifest themselves in people. Some examples of these diseases are brain tumors, Alzheimer's disease, breast cancer, lung cancer, and cardiovascular disease. The ability to easily diagnose and treat these conditions is made possible by medical image processing [1]. Multimodel computing is efficient in Alzheimer disease detection. Multimodel

computing is useful in many medical applications like lung cancer detection, breast cancer detection, and medical image classification and detection.

Alzheimer's disease is now the primary research interest (AD). Alzheimer's disease is an example of a kind of brain disorder. Some of the most common types of medical imaging are positron emission tomography (PET), positron X-ray tomography (X-ray), computed tomography (CT), ultrasound, and magnetic resonance imaging (MRI). The circumstances of the patient are what guide the selection of the appropriate imaging technique. When diagnosing a patient with a bone problem, it is a common practice to take x-rays of the affected area of the patient's body. The diagnosis of Alzheimer's disease is the focus of the research being conducted right now, which makes use of the magnetic resonance image processing [2].

The medical needs of almost every person in this nation are met by the healthcare system provided by the pharmaceutical sector. The great majority of conditions that affect the human health are due to dysfunctions that occur inside the body's organs and tissues. Medical image processing entails developing a visual portrayal of the inside of a body using a range of imaging technologies in order to discover and cure problems [3]. This may be done via the use of computer software. The discipline of medical image processing, which is a subfield of image processing that contributes to the improvement of public health, has a number of challenges.

The magnetic resonance imaging (MRI) technique is a noninvasive medical imaging method that works by producing the images of inside organs, bones, and other human tissues via the use of high magnetic fields and radio waves. The proton magnetic resonance imaging (MRI) scanner makes use of a strong magnetic field in order to arrange the hydrogen atoms' protons within the body. After that, radio waves are used to spin the protons. After the radio waves have been turned off, the protons will realign themselves by producing new radio waves on their own. This piece of equipment is capable of picking up radio waves and producing an image of them. Magnetic resonance imaging, more often known as MRI, is the technique of choice for situations in which high-resolution images are needed, such as when abnormalities of the brain are being diagnosed. Using this strategy, one may be able to prevent exposure to radiation with a high energy level [4].

By altering the order in which the radio waves are received, various images may be produced. The repetition time refers to the amount of time that has passed between the two successive radio wave sequences that have been applied to the same slice (TR). It is possible to differentiate between the various tissues by using their respective relaxation times. Both T1 and T2 relate to relaxation times in the longitudinal and transverse axes, respectively, as well as relaxation durations in spin lattice and spin-spin lattices. Additionally, T1 and T2 refer to relaxation periods in spin-spin lattices. The time constant T1 determines the pace at which the stimulated protons return to their original state of equilibrium. T2 is the most important variable to consider when attempting to compute the rate at which the protons approach

equilibrium or become out of phase with each other. MRI images may also be classified according to their sequences using terms such as T1-weighted and T2-weighted MRI. It is possible to differentiate between T1- and T2-weighted photographs by using the cerebrospinal fluid (CSF) that is found in the brain. The cerebrospinal fluid (CSF) appears black in T1-weighted imaging, but it appears bright in T2-weighted pictures. Figure 1 displays an image that is weighted T1 as well as a photo that is weighted T2 [5].

Dr. Alois Alzheimer was the one who first recognized the symptoms of Alzheimer's disease (AD) in 1906. Each year, more than two million people in World are given a diagnosis of Alzheimer's disease [6]. It is a disease that causes the brain to deteriorate over time. It starts off mild, but over the course of time, it becomes increasingly more severe. Alzheimer's disease causes damage to brain cells as well as the death of those cells. Memory in humans is especially susceptible to this. Memory loss is the first indication of Alzheimer's disease, but as the disease progresses and more brain cells die, additional symptoms arise. These include shifts in mood and behavior, difficulties communicating, and problems remembering the names of known people, places, or recent events. People who have Alzheimer's disease may eventually reach a point when they are unable to do the tasks required of them on their own. They put their whole well-being in the hands of another person. This is due to the fact that brain changes that accompany normal ageing may bring about memory loss. To put a stop to the progression of Alzheimer's disease (AD), however, it is essential to diagnose the condition at an early stage. Even though there is currently no cure for Alzheimer's disease, early discovery of the condition might potentially reduce or stop the progression of the illness.

Alzheimer's disease (AD) is characterized largely by the death of nerve cells and tissues in the brain, which ultimately leads to a reduction in brain volume. The brain is both an essential component of the nervous system as well as one of its most intricate components. The human brain is made up of three different parts: the cerebrum, the brainstem, and the cerebellum. The largest and most complex part of the neurological system, the cerebrum is located at the front of the skull. This part of the brain is involved in a broad range of mental functions, such as memory, reasoning, problem solving, emotional regulation, and the perception of sound and light. There are two hemispheres of brain in the cerebrum: left and right. The grey matter of the cerebral cortex, the outermost layer of the cerebrum, dominates this region. In the brain's cortex, which is made up of several layers, there are billions of nerve cells. "White matter" is a lengthy nerve fibre that connects the different parts of the brain together and is often referred to as such. One of the early warning signs of Alzheimer's disease is a reduction in grey matter in the cerebral cortex. The hippocampus and basal ganglia are only two examples of the many subcortical structures found inside the cerebrum. The four lobes of the cerebrum are the frontal lobe, temporal lobe, occipital lobe, and parietal lobe. As Alzheimer's disease progresses, the temporal lobe of the brain is most often afflicted. Other brain areas will also begin to degenerate as the illness progresses.

(a)   (b)
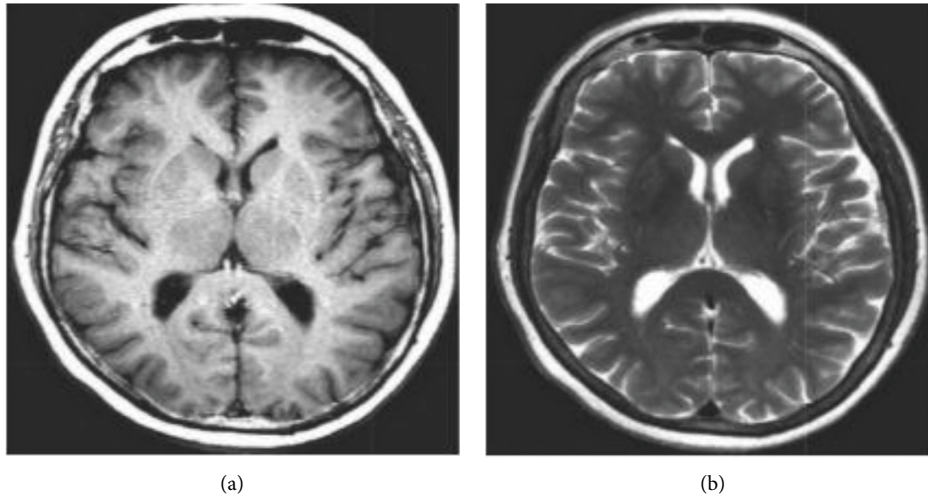
FIGURE 1: (a) T1-weighted MRI, (b) T2-weighted MRI.

Both movement coordination and the maintenance of balance are crucial functions of the cerebellum, which resides immediately below the cerebrum. Several of the body's autonomic processes, including digestion, respiration, heart rate, and temperature, are controlled by the brainstem, which lies underneath the cerebral cortex and directly in front of the cerebellum. The cerebrum, which is positioned at the top of the skull, is the primary target of Alzheimer's disease (AD) [7].

The magnetic resonance imaging (MRI) technique is helpful for providing a dynamic diagnostic of the structure and volume of the brain. In order to make an accurate diagnosis, it is essential to have the capability of recognizing quick changes in the brain utilizing dynamic analysis. Magnetic resonance imaging, sometimes known as an MRI, is a technique that is frequently used for the purpose of precisely diagnosing Alzheimer's disease in its initial stages. Changes in the hippocampus and entorhinal cortex will be seen in the reports generated by MRIs performed on people with Alzheimer's disease. Because of the possibility of mistakes and inaccuracies introduced by the involvement of humans in the process, it is necessary to use more efficient alternatives such as automated systems. MRI features and machine learning algorithms are used to make an automatic diagnosis of Alzheimer's disease.

Literature review section contains a review of modern techniques for Alzheimer disease detection. Methodology section presents machine learning-based multimodel computing for medical imaging for classification and detection of Alzheimer disease. Images are acquired first. MRI images contain noise and contrast problem. Images are preprocessed using CLAHE algorithm. It improves image quality. Images are segmented using the k-means algorithm. This results in the segmentation of images and identification of region of interest. Useful features are extracted using PCA algorithm. Finally, images are classified using machine learning algorithms. Result section contains details related to input data set and results obtained by various machine learning algorithms. Conclusion section contains major contributions of the research article.

1.1. Medical Image Processing. The phases of medical image processing are image preprocessing, image segmentation, and feature extraction and classification.

1.2. Preprocessing. Preprocessing, often known as augmentation, is the first step in the medical image processing workflow [8]. Image enhancement, often known as preprocessing, is the process of improving an image's quality before it is utilized in following processing steps. This allows the image to be put to better use. This method is used to improve the picture quality since it is possible that an inaccurate diagnosis might result from an inaccurate image. Before they can be used in the diagnostic process, medical images often need some kind of update. The removal of noise, improvement of contrast, and "skull stripping" are three of the most common MRI image modifications.

There will always be some level of noise in an image, regardless of the method used to shoot the picture or the gear that was used. The process of reducing unwanted noise from an image is referred to as noise reduction. Noise may be reduced using a number of filtering techniques, many of which are available via digital image processing.

The process of boosting the dynamic range of an image's intensity values is referred to as "contrast enhancement," and it is one of many methods that can be used to improve the contrast of an image. The image's primary components become more distinguishable from the background when the contrast has been increased.

In a brain MRI, it is possible to view tissues that are not part of the brain, such as the skull, skin, fat, muscle, and the neck. The existence of these non-brain tissues makes it difficult to do further study on those tissues. Skull stripping [3] is the process that is utilized to remove non-brain material prior to continuing the examination of the patient.

In this study, skull stripping was accomplished by the use of entropy-based thresholding in conjunction with several morphological methods.

1.3. Segmentation.

In the field of medical image processing, segmentation [9] is a technique that is used to separate the diseased region from the rest of the image. It does this by using aspects like as intensity or texture to separate an image into discrete pieces depending on how similar they are to one another. It is possible to make use of the divided region of interest in order to readily extract crucial information for the purpose of sickness detection.

Using approaches that are based on a threshold, images may be segmented. It is a translation from binary to pictures. In this method, every pixel in the image is either completely black or completely white. In a digital picture, the intensity of a pixel is compared to a threshold, which is a constant value denoted by the letter "T." Based on this value, the pixel is either replaced with a white or a black pixel.

If you use the strategy of region-based segmentation, you may be able to find regions that have certain qualities in common with one another. It does this by slicing the image up into distinct subregions, none of which can share the same characteristics with one another. It is possible to divide it into two groups, namely, the expansion of regions and the division and combination of regions. In the method for growing regions, a seed point is utilized, and the area grows outward from that point by connecting pixels that are near to the seed and have properties with the seed itself. The starting point might be one or many seed points, depending on the situation. Using the Split and Merge method, first some random bits are divided into, and then they may be combined and/or split in an attempt to generate unique regions in the image that are of a similar nature to one another.

Edge-based segmentation software often uses this method, which involves dividing an image into sections depending on the edges or boundaries that it has. The field of image processing provides a wide number of options for finding edges. Methods for locating edges often focus on identifying discontinuities or changes in intensity. The intensity levels of an image item rapidly change as they get closer to the picture's borders.

K-means clustering is the method that is used the majority of the time in the process of medical image segmentation. During the clustering technique, the image is cut up into distinct groups or clusters that do not overlap with one another. In this image, there are a few different clusters, and each of these clusters has its own unique set of reference points to which each pixel is assigned. The k-means clustering method utilizes k reference points and results in k distinct groups.

1.4. Feature Extraction.

In order to facilitate the process of sickness diagnosis, a technique known as feature extraction is used to glean important and pertinent information from a sectioned off region of interest. The recovered characteristics have a direct bearing on the level of accuracy that may be achieved while diagnosing a disease. If the traits that are produced from the classification are employed appropriately, it is possible that more informed judgments may result. When it comes to the processing of medical images, shape and size are the most crucial components. Because Alzheimer's disease changes the size and shape of the brain, the classification results of brain MRIs performed to diagnose Alzheimer's disease may be affected as a result.

1.5. Classification.

Image processing often continues with categorization as the next step after the extraction of features. Both the classifier that was applied to the image in order to classify it and the recovered features from the image itself are directly responsible for the result of the classification. When the class of an image cannot be determined, the classifier will give the picture a label. In this instance, the classification for the class might be either normal or abnormal (with disease). In the field of medical image processing, two of the classifiers that are used most often are the k-nearest neighbor and the support vector machine [10].

For the purpose of data classification, a strategy known as the k-nearest neighbor algorithm (k-NN) is used. In the k-NN classification system, the majority of an image's neighbors decide whether or not an image should be considered normal or abnormal. The image gets filed away in a particular classification according to which of its k immediate neighbors it has the most characteristics with. A positive integer that goes by the name k is one of the numbers that comes up most often.

A binary classifier is a technique that is recognized by the acronym SVM (which stands for support vector machine). It is possible to predict the class of each feature vector by utilizing a feature vector as the input. The approach in question creates two categories, which are referred to as normal and abnormal, and places a significant gap between them. Based on the data that is at hand, it is clear that the SVM classifier generates very good results when it is coupled with an appropriate kernel. The equation for the hyperplane of the linear SVM is as follows:

$$ww.ll + bb = 00, \tag{1}$$

where, b = real number, w = normal vector to the hyperplane, l = feature vector.

Artificial neural networks, also known as ANNs, are often used in the process of classifying medical pictures for the purpose of disease diagnosis. The functioning of the ANN is quite similar to that of the human brain. By looking at a collection of pictures that have already been labelled, it is able to acquire the knowledge necessary to make an accurate estimate about the category that an image belongs to. Artificial neurons, which are the building blocks of an ANN, are designed to mimic their natural counterparts in the human brain. Neurons are connected to one another along their edges. Weights may be assigned to neurons and edges, and these can be altered at any point throughout the learning process. The majority of artificial neural networks are constructed with three layers: an input layer, a hidden layer, and a final layer that is responsible for outputting the signal. It is possible that there is only one hidden layer, that there

are numerous hidden levels, or that there are none at all. Weights are adjusted inside a layer that is concealed from view until the desired result is obtained.

1.6. Literature Survey

1.6.1. Literature Review of Image Preprocessing Algorithms. Priyanka and Balwinder [11] proposed the possibility of using a method known as the Median Filter in order to get rid of salt and pepper noise as well as Poisson noise in photographs. For instance, the output intensity value of the pixel that is to be processed is made by sliding a window across the image. This value is then determined by using the median intensity of the pixels that are included inside the window to calculate the value. In addition, the median filter maintains the borders of a picture while simultaneously minimizing the amount of random noise. The values of every pixel are permanently set to the median of the values of the pixels that are immediately around them. After that, it is put to use to get rid of these noises, and after that, the bounding box approach is carried out to find the tumor.

Yousuf and Nobi [12] state that the research into the creation of order statistics filters has resulted in the invention of a simple solution that is both efficient and effective for cutting down on the amount of noise present in medical images. As can be seen in the above illustration, the median and mean filters are used in order to get the pixel value of a picture that is free of noise. In addition, it may be used to lessen the appearance of visual artefacts such as Rician noise.

Jaya et al. [13] came up with the idea for the weighted median filter. The use of a weighted median filter for denoising allows for the reduction of high frequency components as well as the removal of salt and pepper noise from images without causing the edges to become distorted. Additionally, it may be used to extract each pixel from a window of pixels that is 3 by 3, and then analyze the mean value of the foreground and background pixels, as well as the contrast value. The background noise in a picture may be removed using an anisotropic filter that was created by Ramalakshmi and Chandran [14].

Comparisons were then made between the wavelet denoising and the Gaussian smoothing techniques [15]. In the steps leading up to the reconstruction of MR images, a wavelet-domain Wiener-filtering technique was used [16]. However, because of the underlying wavelet structure, the wavelet-based techniques that are often utilized have the potential to result in considerable artefacts being introduced into the processed images. One of the most prevalent approaches to denoising that is now accessible is called a maximum posteriori estimate methodology. Rician noise is taken into consideration in these approaches by the use of both a data probability term and a spatial smoothing prior [17]. Empirical Bayes was used by Awate and Whitaker [18] for the MRI denoising process. The Markov Probability Density Function (PDF), which is used as a prior in Bayesian estimation, is used to do an analysis of the distortions that are present in the data.

1.6.2. Literature Review of Feature Extraction Algorithms. Either image segmentation or the registration of a brain Atlas over the image may be used to count the voxel values in important anatomical locations. This can be done either manually or automatically. In spite of these challenges, the structural parcellation of the brain may not be able to adjust to the effects of the illness.

To begin, Khajehnejad et al. [19] employed voxel morphometry analysis to extract from actual MRI volumes and Gray Matter (GM) segmentation volumes some of the most likely AD-relevant elements of brain imagery. These volumes were segmented based on grey matter. The characteristics that set a healthy brain from one afflicted by Alzheimer's disease must be included into the characteristics of the features. After that, a dimension reduction using Principal Component Analysis (PCA) is carried out on the collected features in order to conduct an analysis that is not only quicker but also more accurate. In order to make use of the returned features, a hybrid manifold learning framework has been proposed here. This framework brings feature vectors into a subspace.

Assessment of the cortical thickness that is quick, accurate, and completely automated has been created by Querbes et al. [20]. There may be a connection between the existence of histopathological validated anomalies and the progression of cortical atrophy, which is evaluated by the cortical thickness. It is possible to make adjustments to the volume in this way by utilizing the estimated total brain volume by Cuingnet et al. [21]. Cortical thickness testing provides the chance for results that are less dependent on the operator, in contrast to hippocampal volume measurement, which is highly dependent on the individual doing the test by Higdon et al. [22]. Those with higher levels of education and more severe brain damage have a greater propensity to conceal indications of dementia due to their cognitive reserve. This may be perplexing to people who are not acquainted with the condition since it occurs less often in people with a higher levels of education.

Using ROI-based techniques, one or more essential components of the brain, such as the cingulum, the corpus callosum, and similar structures, may be characterized. Studies have indicated that the development of neurodegeneration in Alzheimer's disease (AD) affects the regions of the brain that are situated in the limbic and neocortical areas, as well as the temporal and temporal lobes of the brain. Additionally, the temporal and temporal lobes of the brain are affected. The atrophy of the medial temporal region, and in particular the atrophy of the hippocampi, is generally recognized as a sensitive biomarker of Alzheimer's disease (AD) [23]. As a consequence of this, hippocampi have been employed in a number of studies as a biomarker for early-onset Alzheimer's disease.

In addition, Sakthivel et al. [24] incorporated not just the information that can be found in text and photographs, but also the direct input from the doctor. It is possible for a feature to have coefficients that match up to an image spectral transform, such as Fourier or Discrete Cosine Transform (DCT) coefficients, statistics on picture gradients, and other such things. Two characteristics that are utilized to

characterize the brain images are called Local Binary Patterns (LBPs) and Discrete Cosine Transforms (DCT).

Researchers have successfully extracted three separate features from the MR scans of the brain for the very first time by combining the grey matter volume, the Gray-Level Co-Occurrence Matrix (GLCM), and the Gabor feature. This achievement marks a first. The results of the experiments indicate that a greater performance may be accomplished by the multifeature fusion of these characteristics, which may gather both 2D and 3D information on brains. This can be accomplished by combining the features of many brain scans. The researchers Agarwal and Mostafa [25] employed visual image similarity as a tool to assist in the early identification of Alzheimer's disease. It displays how well the brain images may be categorized depending on the information provided by the user. Calculations of Circular Harmonic Functions and Scale Invariant Features Transform (SIFT) descriptors are performed close to the hippocampus, same as in a prior work [26]. After that, a number of classification schemes are used in order to make comparisons between the photographs.

1.6.3. Literature Review of Classification Methods.

Because there are so many voxels in the brain, the qualities that may be deduced from the combination of voxels are quite specific and precise. LDA is a well-known method for reducing the dimensions of a problem; another name for this method is the Fisher linear discriminant (FLD). As an example, linear discriminant analysis (LDA) makes use of a linear discriminant function to locate low-dimensional linear combinations of variables that provide the most accurate description of the data. In order to do this, the between-class scatter matrix is made larger while the within-class scatter matrix is made smaller [27].

A technique based on machine learning that was developed by Long et al. can differentiate between patients suffering from moderate cognitive impairment (MCI) and healthy older persons. It is also possible to use this method to forecast whether or not a patient diagnosed with MCI would eventually develop Alzheimer's disease (AD). After this phase, a symmetric diffeomorphic registration, an embedding approach, and a learning method for determining the distance between the subjects are available. These results were obtained when the amygdala and/or hippocampus were used as the area of interest (ROI): 96.5 percent for mild AD identification, 91.74 percent for progressive MCI differentiation, and 88.99 percent for classification of the two types of MCI. By using the macroscopically distinct shapes that occur in each pair, this technique has maximized its differentiation potential.

Zhao et al. [28] invented the Iterative Trace Ratio (iITR) to address the TR-LDA (Trace Ratio Linear Discriminant Analysis) issue for dementia diagnosis. iITR outperformed PCA, LPP, and the Maximum Margin Criterion in terms of outcomes. [Reference required] (MMC). Image features used to distinguish between AD and FTD in LDA were compressed using the Partial Least Squares (PLS) approach by Horn et al. [29]. The accuracy, sensitivity, and specificity of the SPECT pictures obtained by the researchers were all over 84%.

Those classifiers that come from the Nave Bayes theorem are in the same family as other probabilistic classifiers since they make the assumption of feature independence. Decision models for Alzheimer's disease (AD), moderate cognitive impairment (MCI), and neurological dementia (NC) were developed by Seixas et al. [30]. In the end, they concluded that the Bayesian network decision model outperformed several well-known classifiers, including the naive Bayes, the logistic regression model, the multilayer perceptron ANN, and the decision table. Liu and Shen [31] multifold Bayesian Kernelization approach is better at distinguishing between Alzheimer's disease (AD) and non-converter (NC) MCI, but it is less accurate at identifying MCI-converter (MCIc) and non-converter (MCIn).

SVMs allow for the construction of hyperplanes with high or indeterminate dimensions, which may then be utilized for a variety of applications, including classification and regression. Because SVMs have a lower generalization error compared to other classifiers, they are frequently used to solve pattern-classification problems with limited sample sizes [32]. There are a total of 120 subjects, with 40 ADs, 40 MCIs, and 40 NCs allocated to each of the three categories, respectively. In the beginning, each subject was subjected to filtering and normalization, and after that, K-Nearest Neighbor (KNN) or Support Vector Machine was used in order to extract a total of twelve features (SVM). It was determined that several permutations and combinations of each feature should be tried in order to uncover the best characteristics for categorizing the data. This was done with the intention of finding out which ones were the most accurate. For a random selection of test data using SVM and KNN, the results showed an accuracy of 95.833 percent on average, with SVM polynomial order three yielding the highest accuracy at 97.92 percent, and KNN with $K = 6$ and $K = 7$ yielding the lowest accuracy at 95.83 percent. It was shown that there was a high level of accuracy in classification across all three clinical groupings. The Master Characteristics of the images were extracted with the use of a quick discrete wavelet transform (DWT), and then the Principal Component Analysis (PCA) was utilized to conduct additional research on the distinguishing characteristics that were discovered (PCA). There are a total of five distinct decision models that are each given a unique subset of the key feature vectors. Models such as the J48 decision tree and KNN, Random Forest (RF), and LS-SVM with polynomials and radial basis kernels are used as part of the classification models.

The covariance method was used to study several feature correlation technologies and enhance the SVM-RFE algorithm by means of the covariance technique. This was accomplished via the usage of the covariance approach. The recently devised strategy seems to be beneficial, based on the results of analyses conducted on the publicly available ADNI database. It also suggests that using a combination of numerous features is preferable than making use of a single trait on its own.

It is possible to estimate for approximation functions that are dependent on a large number of unknown inputs by using models that are based on artificial neural networks (ANNs), which are a subset of models that are impacted by biological neural networks. They have taken the place of rule-based programming as the go-to solution for a broad variety of complicated issues, and with good reason. Standard discriminant function analysis was shown to have lower sensitivity and accuracy compared to employing an artificial neural network (ANN) for MRI-based dementia classification.

Luo et al. [33] presented a deep learning system that was built on 3D brain MRI as a way for automatically identifying the Alzheimer's disease. This system was stated as being able to do so. Convolutional Neural Network (CNN) is used in order to diagnose Alzheimer's disease (AD). One of the most important distinctions is that in the identification of AD, the three-dimensional structure of the brain is regarded to be complete, making it possible to make an accurate diagnosis. The Convolutional Neural Network (CNN) that was used in this investigation contains three consecutive sets of processing layers, two levels that are totally connected, and a classification layer. Each of the three different groups has a structure that consists of a convolutional layer, a pooling layer, and a normalizing layer. All three layers are included.

2. Methodology

This section presents machine learning-based multimodel computing for medical imaging for classification and detection of Alzheimer disease. Images are acquired first. MRI images contain noise and contrast problem. Images are preprocessed using CLAHE algorithm. It improves image quality. Images are segmented using the k-means algorithm. This results in segmentation of images and identification of region of interest. Useful features are extracted using PCA algorithm. Finally, images are classified using machine learning algorithms. Block diagram of the model is shown in Figure 2.

In order for an image to be accurately recognized, the process of background extraction has to be able to adapt to the one-of-a-kind characteristics of the particular photograph being used. Within CLAHE, the histogram is only constructed for the pixels that are immediately next to it. By imposing a "clip level" on the height of the local histogram and, therefore, the maximum contrast enhancement factor, CLAHE limits the amount of contrast alteration that may be performed. Because of this, there is noticeably less noise in the final image. CLAHE is better to other methods in its capacity to enhance the look of mammography in minute places. A white background makes the lesions more obvious to the naked eye. In spite of the fact that this method makes it simpler to differentiate between signal and noise, the images still include a significant amount of graininess [34].

Segmentation is a method that is employed in the area of medical image processing. Its purpose is to separate the diseased portion of an image from the healthy parts of the picture. It accomplishes this goal by segmenting a picture into distinct parts based on how closely those pieces
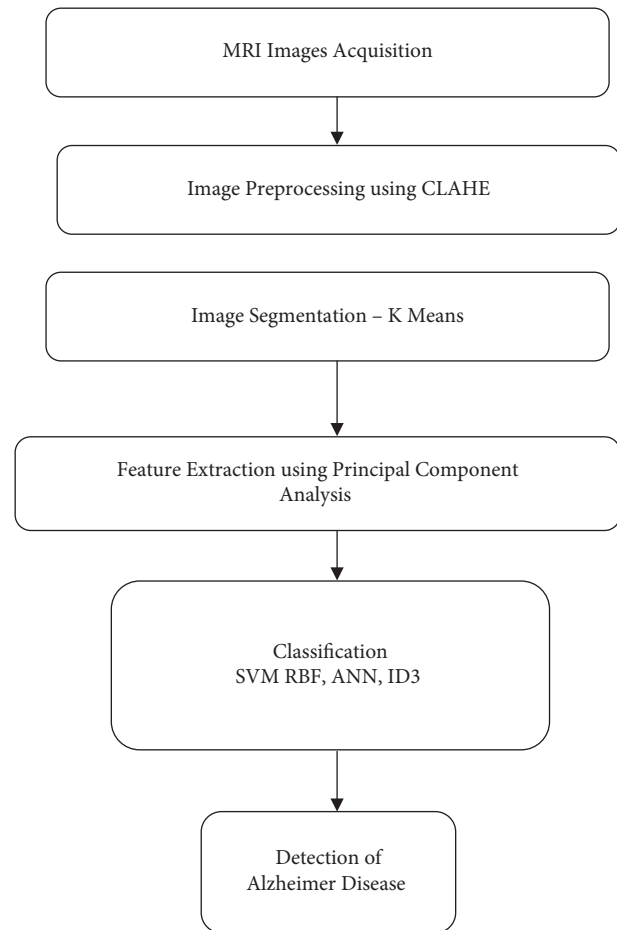
Figure 2: Machine learning-based multimodel computing for medical imaging for classification and detection of Alzheimer disease.

resemble one another. This is done using image characteristics such as intensity and texture. It is feasible to make use of the segmented area of interest in order to quickly extract important information for the aim of diagnosing a disease. In the process of medical picture segmentation, the technique known as k-means clustering is the one that is used the vast majority of the time. During the clustering process, the picture is divided into a number of unique groups, also known as clusters, which do not overlap with one another. This picture has a few distinct clusters, and each of these clusters has its own one-of-a-kind set of reference points to which each pixel is allotted. The k-means clustering technique divides the data into k different groups using the k reference points that are provided in the process [35].

The approach known as Principal component analysis (PCA) is used in the process of feature extraction [36]. The method of principal component analysis (PCA) to reducing linear dimensions might be helpful in data analysis and compression [37]. Using this method, which involves finding orthogonal linear combinations of the attributes of the initial data set, it is possible to combine qualities that are not connected with one another.

A binary classifier is a technique that is recognized by the acronym SVM (which stands for support vector machine). It is possible to predict the class of each feature vector by utilizing a feature vector as the input. The approach in question creates two categories, which are referred to as normal and abnormal, and places a significant gap between them. Based on the data that is at hand, it is clear that the SVM classifier generates very good results when it is coupled with an appropriate kernel [37]. SVM works better with RBF function. The equation for the hyperplane of the linear SVM is as follows:

$$ww.ll + bb = 00, \tag{2}$$

where, b = real number, w = normal vector to the hyperplane, l = feature vector.

The process of categorizing medical images for the goal of illness detection often makes use of artificial neural networks, which are sometimes referred to by their acronym, ANNs. The way in which the ANN works is quite comparable to the way in which the human brain does. It is possible to learn the information essential to make an accurate guess about the category that an image belongs to by looking at a collection of photographs that have previously been labelled. This collection of pictures has already been classified. Artificial neurons, which are the fundamental components of an ANN, are intended to function in a manner similar to that of their natural counterparts in the human brain. Along the margins of their bodies, neurons are linked to one another. It is possible to provide weights to neurons and edges, and these weights may be changed at any time in the course of the learning process. The majority of artificial neural networks are built with three layers: an input layer, a hidden layer, and a final layer that is in charge of outputting the signal. The majority of artificial neural networks are built with three layers: an input layer, a hidden layer, and a final layer. It is conceivable that there is only one hidden layer, that there are several hidden layers, or that there are none at all. All of these outcomes are feasible. Adjustments are made to the weights that are contained inside a layer that is hidden from view until the outcome that is sought is achieved [38].

J. Ross Quinlan is the one who came up with the ID-3 technique, which is also referred to as the Iterative otomiser-3. This was the first strategy to use a dynamic decision tree as its foundation. This technique utilises a measure of information gain in addition to entropy as its primary metric. Beginning with a nodule is the first step in an iterative process that establishes an entropy value for each of the functional qualities. Under the strictest sense, data sets that have the lowest error rates are referred to be "split attributes" in the theories of entropy and "information gain" (entropy). Since there is no definitive categorization of the target classes, the algorithm must repeat through its own stages for each individual subset of data. The nonterminal nodes that make up a branch's terminal nodes are what are referred to as that branch's terminal nodes. The split attribute may be utilized to determine whether nodes inside a tree structure are not terminal since these nodes do exist [39].

3. Result, Analysis, and Discussion

This experiment makes use of the data gathering that was performed by OASIS [40]. In all, this dataset has 416 different samples. Machine learning strategies such as SVM-RBF, ANN, and ID3 are used while classifying the data. All forms of Alzheimer's disease, including mild Alzheimer's disease, huntington disease, and even normal MRI scans, are grouped together under the umbrella term Alzheimer's disease. The fifty pictures that make up each category were selected at random from a pool of two hundred images.

Five parameters such as accuracy, sensitivity, specificity, precision, and recall are used in this study to compare the performance of different algorithms.

$$
\begin{aligned}
\text{Accuracy} &= \frac{(TP + TN)}{(TP + TN + FP + FN)}, \\[6pt]
\text{Sensitivity} &= \frac{TP}{(TP + FN)}, \\[6pt]
\text{Specificity} &= \frac{TN}{(TN + FP)}, \\[6pt]
\text{Precision} &= \frac{TP}{(TP + FP)}, \\[6pt]
\text{Recall} &= \frac{TP}{(TP + FN)},
\end{aligned}
\tag{3}
$$

where, TP = True Positive. TN = True Negative. FP = False Positive. FN = False Negative.

As can be shown in Figures 3–7, the SVM-RBF classifier offers the highest level of accuracy out of all of the available options for diagnosing Alzheimer's disease. ANN and Random ID3 come up at second and third place, respectively, when it comes to sensitivity, specificity, accuracy, and recall. The ANN algorithm's sensitivity and recall are much higher than those of the other classifiers. SVM-RBF is superior to the other classifiers in terms of its level of specificity.
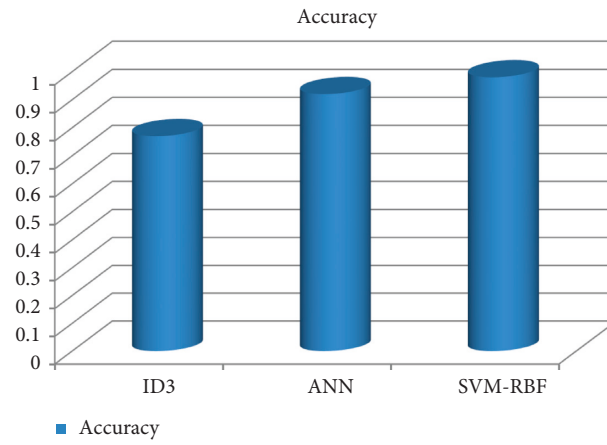
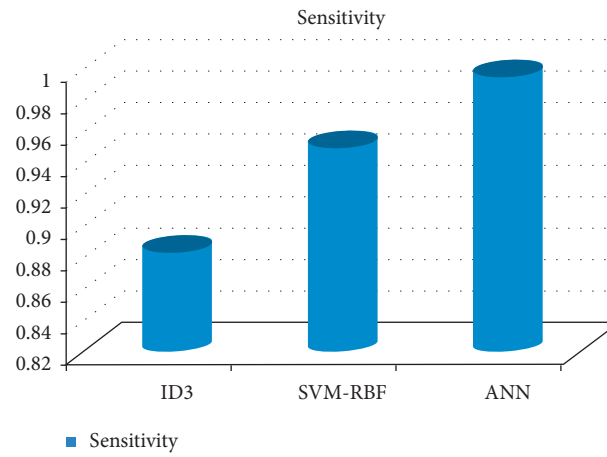Figure 3: Accuracy comparison of machine learning algorithms.

Figure 4: Sensitivity comparison of machine learning algorithms.
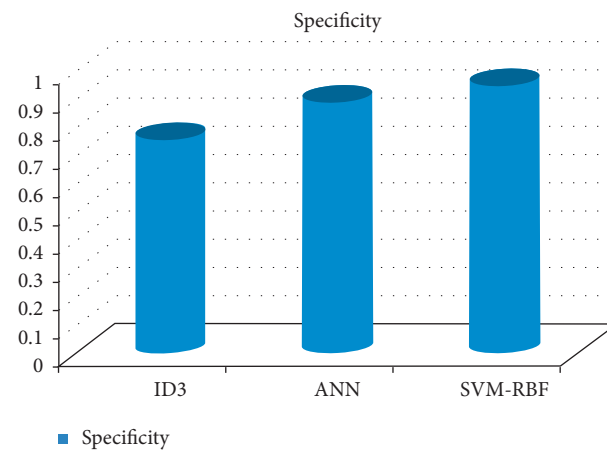

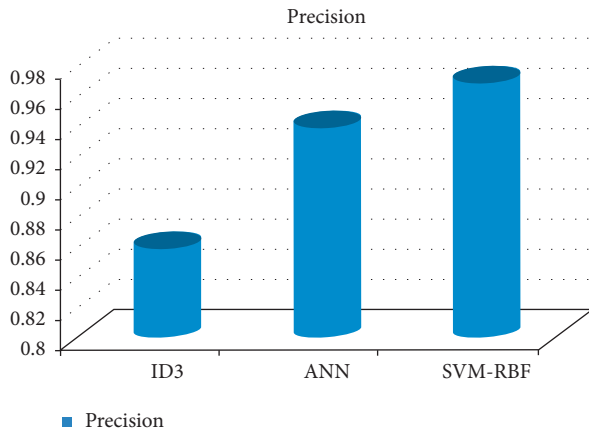Figure 5: Specificity comparison of machine learning algorithms.

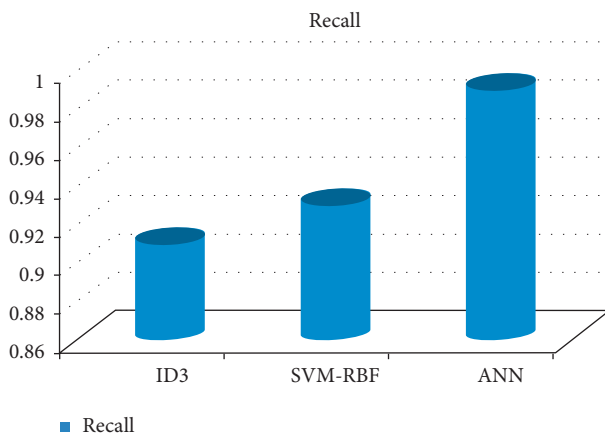Figure 6: Precision comparison of machine learning classifiers.

Figure 7: Recall comparison of machine learning classifiers.

4. Conclusion

Alzheimer's disease is a progressive deterioration of brain function that happens over time. It begins off not being very serious, but as time passes, it quickly escalates into a much more serious condition. Alzheimer's disease is a degenerative neurological condition that leads to both the damage and death of brain cells. Memory in humans is particularly prone to being affected by this. The initial sign of Alzheimer's disease is memory loss, but as the illness advances and more brain cells die, more symptoms emerge. Alzheimer's disease is characterized by a progressive loss of brain cells. In the field of medicine, "medical image processing" refers to the process of creating a visual representation of the internal workings of a body by using a variety of imaging technologies in order to diagnose and treat illnesses. In this study, we describe a machine learning-based multimodel computing approach for medical imaging, with the goals of classifying patients and locating Alzheimer's disease. The process begins by acquiring images. MRI imaging has noise and contrast issue. The CLAHE algorithm is used as a preprocessor for the images. It results in a higher overall picture quality. The k-means technique is used in order to separate the images. This leads to the segmentation of the pictures as well as the detection of the

area of interest. Utilizing the PCA technique, useful characteristics are retrieved. In the last step, photos are categorized with the help of machine learning algorithms. SVM-RBF classifier offers the highest level of accuracy out of all of the available options for diagnosing Alzheimer's disease. ANN and Random ID3 come up at second and third place, respectively, when it comes to sensitivity, specificity, accuracy, and recall. The ANN algorithm's sensitivity and recall are much higher than those of the other classifiers. SVM-RBF is superior to the other classifiers in terms of its level of specificity.

Data Availability

The data used to support the findings of the study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] R. F. Buckley, A. P. Schultz, T. Hedden et al., "Functional network integrity presages cognitive decline in preclinical Alzheimer disease," *Neurology*, vol. 89, no. 1, pp. 29–37, 2017.

[2] K. N. H. Dillen, H. I. L. Jacobs, J. Kukolja et al., "Functional disintegration of the default mode network in prodromal Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 59, no. 1, pp. 169–187, 2017.

[3] M. Shabaz and U. Garg, "Predicting future diseases based on existing health status using link prediction," *World Journal of Engineering*, vol. 19, no. 1, pp. 29–32, 2021.

[4] A. Sharma, S. Kaur, N. Memon, A. Jainul Fathima, S. Ray, and M. W. Bhatt, "Alzheimer's patients detection using support vector machine (SVM) with quantitative analysis," in *Neuroscience Informatics* vol. 1, no. Issue 3, , p. 100012, Elsevier BV, 2021.

[5] S. Chaudhury, N. Shelke, K. Sau, B. Prasanalakshmi, and M. Shabaz, "A novel approach to classifying breast cancer histopathology biopsy images using bilateral knowledge distillation and label smoothing regularization," in *Computational and Mathematical Methods in Medicine*, D. Koundal, Ed., vol. 2021pp. 1–11, Article ID 4019358, 2021.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[7] S. Chaudhury, A. N. Krishna, S. Gupta et al., "Effective image processing and segmentation-based machine learning techniques for diagnosis of breast cancer," *Computational and Mathematical Methods in Medicine*, vol. 1, Article ID 6841334, 2022.

[8] J. Godara, R. Aron, and M. Shabaz, "Sentiment analysis and sarcasm detection from social network to train health-care professionals," *World Journal of Engineering*, vol. 19, no. 1, pp. 124–133, 2021.

[9] A. Gupta and N. Koul, "SWAN: a swarm intelligence based framework for network management of IP networks," in *Proceedings of the International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, Sivakasi, India, December 2007.

[10] A. S. Zamani, L. Anand, K. P. Rane et al., "Performance of machine learning and image processing in plant leaf disease

detection," *Journal of Food Quality*, vol. 2022, p. 1, Article ID 1598796, 2022.

[11] B Priyanka and S. Balwinder, "An improvement in brain tumor detection using segmentation and bounding box," *International Journal of Computer Science and Mobile Computing*, vol. 2, pp. 239–246, 2013.

[12] M. A. Yousuf and M. N. Nobi, "A new method to remove noise in magnetic resonance and ultrasound images," *Journal of Scientific Research*, vol. 3, no. 1, pp. 81–89, 2010.

[13] J. Jaya, K. Thanushkodi, and M. Karnan, "Tracking algorithm for denoising of MR brain images," *International Journal of Computer Science and Network Security*, vol. 9, pp. 262–267, 2009.

[14] C. Ramalakshmi and A. J. Chandran, "Automatic brain tumor detection in MR images using neural network based classification," *Biometrics and Bioinformatics*, vol. 5, no. 6, pp. 221–225, 2013.

[15] A. M. Wink and J. B. Roerdink, "Denoising functional MR images: a comparison of wavelet denoising and Gaussian smoothing," *IEEE Transactions on Medical Imaging*, vol. 23, no. 3, pp. 374–387, 2004.

[16] R. Wirestam, A. Bibic, J. Lätt, S. Brockstedt, and F. Ståhlberg, "Denoising of complex MRI data by wavelet-domain filtering: application to high bvalue diffusion-weighted imaging," *Magnetic Resonance in Medicine*, vol. 56, no. 5, pp. 1114–1120, 2006.

[17] S. Basu, T. Fletcher, and R. Whitaker, "Rician noise removal in diffusion tensor MRI," *MICCAI Rician noise removal in diffusion tensor MRI*, vol. 9, no. 1, pp. 117–125, 2006.

[18] S. P. Awate and R. T. Whitaker, "Feature-preserving MRI denoising: a nonparametric empirical Bayes approach," *IEEE Transactions on Medical Imaging*, vol. 26, no. 9, pp. 1242–1255, 2007.

[19] M. Khajehnejad, F. H. Saatlou, and H. Mohammadzade, "Alzheimer's disease early diagnosis using manifold-based semi-supervised learning," *Brain Sciences*, vol. 7, no. 12, p. 109, 2017.

[20] O. Querbes, F. Aubry, J. Pariente et al., "Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve," *Brain*, vol. 132, no. 8, pp. 2036–2047, 2009.

[21] R. Cuingnet, E. Gerardin, J. Tessieras et al., "Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database," *NeuroImage*, vol. 56, no. 2, pp. 766–781, 2011.

[22] R. Higdon, N. L. Foster, R. A. Koeppe et al., "A comparison of classification methods for differentiating fron to temporal dementia from Alzheimer's disease using FDG-PET imaging," *Statistics in Medicine*, vol. 23, no. 2, pp. 315–326, 2004.

[23] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nature Reviews Neurology*, vol. 6, no. 2, pp. 67–77, 2010.

[24] K. Sakthivel, A. Jayanthiladevi, and C. Kavitha, "Automatic detection of lung cancer nodules by employing intelligent fuzzy c means and support vector machine," *Biomedical Research*, vol. 27, pp. 123–127, 2016.

[25] M. Agarwal and J. Mostafa, "Image retrieval for alzheimer's disease detection MCBR-CDS," in *Proceedings of the First MICCAI international conference on Medical Content-Based Retrieval for Clinical Decision Support*, pp. 49–60, Springer-Verlag, Berlin, Heidelberg, 2010.

[26] A. Rueda, J. Arevalo, A. Cruz, E. Romero, and F. A. González, "Bag of features for automatic classification of Alzheimer's disease in magnetic resonance images," in *Iberoamerican Congress on Pattern Recognition*, pp. 559–566, Springer, Berlin, Heidelberg, 2012.

[27] G. Wang and L. Guo, "A novel hybrid bat algorithm with harmony search for global numerical optimization," *Journal of Applied Mathematics*, vol. 2013, pp. 1–21, Article ID 696491, 2013.

[28] M. Zhao, R. H. M. Chan, P. Tang, T. W. S. Chow, and S. W. H. Wong, "Trace ratio linear discriminant analysis for medical diagnosis: a case study of dementia," *IEEE Signal Processing Letters*, vol. 20, no. 5, pp. 431–434, 2013.

[29] J. F. Horn, M. O. Habert, A. Kas et al., "Differential automatic diagnosis between Alzheimer's disease and frontotemporal dementia based on perfusion SPECT images," *Artificial Intelligence in Medicine*, vol. 47, no. 2, pp. 147–158, 2009.

[30] F. L. Seixas, B. Zadrozny, J. Laks, A. Conci, and D. C. Muchaluat Saade, "A Bayesian network decision model for supporting the diagnosis of dementia, Alzheimer's disease and mild cognitive impairment," *Computers in Biology and Medicine*, vol. 51, pp. 140–158, 2014.

[31] F. Liu and C. Shen, "Learning Deep Convolutional Features for MRI Based Alzheimer's Disease Classification," 2014, https://arxiv.org/abs/1404.3366.

[32] S. Liu, W. Cai, Y. Song et al., "Localized sparse code gradient in alzheimer's disease staging," in *Proceedings of the IEEE 35th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC)*, pp. 5398–5401, Osaka, Japan, July 2013.

[33] S. Luo, X. Li, and J. Li, "Automatic alzheimer's disease recognition from MRI data using deep learning method," *Journal of Applied Mathematics and Physics*, vol. 05, no. 09, pp. 1892–1898, 2017.

[34] E. D. Pisano, E. B. Cole, B. M. Hemminger et al., "Image processing algorithms for digital mammography: a pictorial essay," *RadioGraphics*, vol. 20, no. 5, pp. 1479–1491, 2000.

[35] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: a comprehensive survey and performance evaluation," *Electronics*, vol. 9, no. 8, p. 1295, 2020.

[36] S. R. Qureshi and A. Gupta, "Towards efficient Big Data and data analytics: a review," in *Proceedings of the 2014 Conference on IT in Business, Industry and Government (CSIBIG)*, Indore, India, March 2014.

[37] J. Sun, Y. Yang, Y. Wang, L. Wang, X. Song, and X. Zhao, "Survival risk prediction of esophageal cancer based on self-organizing maps clustering and support vector machine ensembles," *IEEE Access*, vol. 8, pp. 131449–131460, 2020.

[38] M. A. Khan and F. Algarni, "A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO-ANFIS," *IEEE Access*, vol. 8, pp. 122259–122269, 2020.

[39] M. A. Mohammed, K. H. Abdulkareem, A. S. Al-Waisy et al., "Benchmarking methodology for selection of optimal COVID-19 diagnostic model based on entropy and TOPSIS methods," *IEEE Access*, vol. 8, pp. 99115–99131, 2020.

[40] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007.

Hindawi

Retraction

Retracted: The Impact of Education Based on New Internet Media Technology on College Students' Mental Health and Biomedical Diagnosis

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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Wang, W. Wei, and J. Zhao, "The Impact of Education Based on New Internet Media Technology on College Students' Mental Health and Biomedical Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3617938, 7 pages, 2022.

Research Article

The Impact of Education Based on New Internet Media Technology on College Students' Mental Health and Biomedical Diagnosis

Dongmei Wang [iD],[1] Wei Wei,[1] and Jinxue Zhao[2]

[1]*The First Clinical Medical College, Hubei University of Medicine, Shiyan 442000, China*
[2]*Laboratory Management Division, Hanjiang Normal University, Shiyan 442000, China*

Correspondence should be addressed to Dongmei Wang; 19402472@masu.edu.cn

There has been an upsurge in signs of gloom, tension, dietary problems, and other dysfunctional behaviors in undergrad populaces lately. At the same time, the need for advisory services is constantly increasing. Some have interpreted these patterns as mental health emergencies that require immediate investigation and the development of possible treatments to meet the needs of students. Later, other studies have linked the observed increase in side effects to shape individual shape enhancement, especially the widespread use of web-based entertainment, and the time spent on such development is clearly a decrease in psychological well-being. Showed to be related while the use of personalized computing innovations has drastically changed the scene in which undergrads interact with one another and appears to have a significant impact on emotional wellness. Similar advances also offer various opportunities for psychological well-being improvement and dysfunctional behaviour treatment. In this segment, we examine the hardships and open doors for undergrad psychological wellness that PC gadgets give. We accentuate potential for extra examination in this field, as well as ways for people and associations to draw in more benefits with these advances in valuable and health-advancing ways.

1. Introduction

When it comes to mental health and health practices, the college years are crucial. Almost 70% of Americans sign-up for school not long after secondary school [1], and about 3/4 of all lifetime instances of mental issues start before the age of 24 [2]. There has been an upsurge in revealed side effects of mental sicknesses in understudied populations as of late. A major epidemiological investigation has discovered that psychological well-being analysis expanded from 22% to 36% among undergrad respondents during the last ten years [3]. In excess of 95% of school prompting center bosses said that the number of students with major mental issues was creating stress on their grounds in a survey [4–5]. The nation over, a rising number of understudies seems, by all accounts, to be in trouble. Some have alluded to the announced ascent

in mental sickness side effects and interest in administrations as a "grounds emotional well-being emergency" [6–8]. Many individuals have started to theorize on the reasons for this supposed emergency. A piece of these increments could be credited to psychological well-being support given before school, which has helped understudies in acquiring admission to school [9], as well as decreases in emotional well-being shame, which might bring about understudies being more willing to disclose and seek help for psychological wellness issues [10–12].

Individualized computing innovations, for example, cell phones, present various issues and potential for undergrad psychological well-being. We are living in a time of extraordinary social networking and instructive asset accessibility. Some might guarantee that we are living in an unrivaled time of social prohibition mindfulness and data

overload. This is confirmed by the production of "feelings of dread toward passing up on a major opportunity," or FoMO, which was at first depicted by an advertising tactician [13–15] and has hence been the subject of numerous exact investigations on psychological wellness and online entertainment use [16–18].

Some information recommends, for instance, that Facebook use is simply dangerous to emotional well-being at the point at which it includes inactive watching of others' posts [19–23] as opposed to more dynamic commitment in friendly associations (see Table 1 for an audit of these discoveries).

It is clear that the expanded use of individualized computing gadgets has altogether altered the scene in which understudies connect with each other. While abuse seems to meaningfully affect psychological well-being, similar advancements offer an assortment of choices for working on emotional well-being and treating mental illness. Balance is fundamental in pretty much every propensity. It would be silly to assume that cell phone and virtual entertainment use are naturally destructive, in light of the fact that they go about as conductors for individuals to interact with their current social encouraging groups of people and channels for individuals to shape new friendly encouraging groups of people. For sure, virtual entertainment stages like Facebook and Instagram are progressively being seen as spots for individual revelations as well as the establishment and maintenance of social connections.

1.1. College Students' Experience of Technology and Social Media. Technological and social media advancements proceed apace, with little comprehension of the broader influence on our well-being. Because of the rapid progress of technology, the majority of the current writing is essentially a couple of years old; the writing is presently not current or valuable. For instance, Apple Inc. released its original iPhone in 2007. Somewhere in the range of 2007 and 2014, there were eight significant models of iPhones (Apple Inc. History, 2015). That compares to roughly one update consistently. Each update brings more data, associations, and general innovative headways [24–27].

1.2. Digital Natives and Internet Abuse. What can you expect to see on a normal college campus? It is usual to see students hooked into some type of electronics, whether on the bus or in line in dining halls. In her book Alone Together, Sherry Turkey describes the ascent of innovative use (2011). As per Turkle (2011), the expected objective of innovation is to associate with others through different means. In any case, as per a few investigations, we have never been more isolated [28–31]. Digital Natives are the individuals who cannot remember a period before intuitive digital media, for example, PCs, cells, and, specifically, PDAs with Internet availability [32]. Computerized natives make up most of the present undergrad understudy populaces. As per Otey, Digital natives experience issues recognizing themselves on the web and disconnected lives (2013). Born between 1980 and 2000, this human era is known as millennials. Since

innovation has generally been a current piece of every thousand years' life, it is oftentimes viewed as a "fundamental and pivotal extremity" [33, 34]. The present understudies are the substance of Digital natives, getting a few advantages from their use of advancements, especially web-based entertainment.

Morgan noticed that the Internet and innovation permitted them to keep in contact with family members and companions back home. Understudies speak with their folks 13 times each week by and large [35]. Specialists and understudies concur that without innovation, understudies may not convey such a strong sensation of help and security due to the straightforward affiliation with home [37]. The accessibility of data basically anywhere helps understudy efficiency and further develops the learning system [36]. As far as educational devices, teachers are presently getting on and empowering understudies to message them late at night. Innovation is utilized to provide assistance, connect people, and promote education.

2. Literature Review

It reviews that today's college students have grown up in a society where virtual entertainment is the standard. Since the regular understudy is a digital native who has never known a period before innovation, utilizing online entertainment and innovation in guidance appears to be legit because of its inescapability in understudies' life. As an instructive apparatus, virtual entertainment and other innovation enhance the growth opportunity by permitting understudies and educators to trade thoughts, encourage coordinated effort and conversation, and draw in and interact through arising social stages [37]. Lately, there has been an increase in web dependence and misuse. Scientists gauge that up to 25% of the overall U.S. populace is dependent on the Internet [38]. Youthful directed a writing study on Internet maltreatment among undergrads. Youthful anticipated the contributing components of Internet maltreatment among undergrads in view of her survey of accessible writing. Web access is limitless and free. Libraries, PC labs, free Wi-Fi, and other such resources work on it to keep a predictable Internet affiliation. Students also have long stretches of unstructured time, which is comparable to a large portion of their unstructured time. To the extent that it has improved, school is habitually the student's first time away from parental as shown in Figures 1 and 2. This newfound versatility could increase the student's yearning to contribute a huge amount of energy to the Internet. Without the parent's or another screen's cautious eye, the understudy is bound to invest energy uncensoredly investigating the Internet. Youthful likewise perceived a component of social terrorizing and distance in Internet misuse. With the rising size of school grounds, one in every five undergrads began utilizing a PC between the ages of 5 and 7. (Flatt, 2013). As a rule, dangerous Internet use proceeds in a cycle and ought to be considered a maladaptive survival technique for individuals who as of now show these qualities. Sadly, for understudies looking for comfort from

TABLE 1: Review of findings.

Negative effects	Positive effects	Enabled interventions
(i) Fear of missing out	(i) Active engagement with peers	(i) Online support groups and message boards
(ii) Peer hyperconnectivity	(ii) Expanded social networks	(ii) Module-based web interventions (e.g., MoodGYNI, beating the blues)
(iii) Companion comparison	Venues for personal disclosures	Skill building apps for resilience, coping skills, and mindfulness (headspace, pacifica, etc.)
(iv) Decreased face-to-face social interaction	Peers can serve as "gatekeepers"	
(v) Impaired development of social skills	(iii) Access to mental health intervention programs	
(vi) Decreased suppression of antisocial behavior		
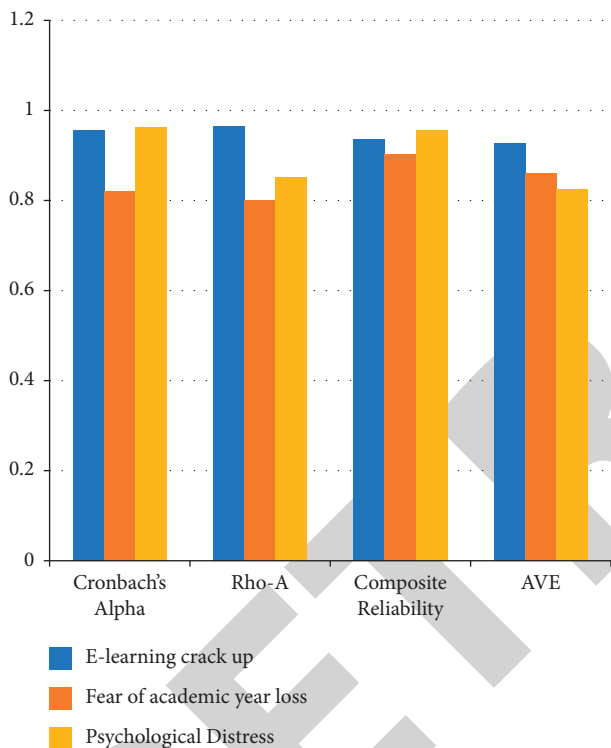
FIGURE 1: Validity analysis.
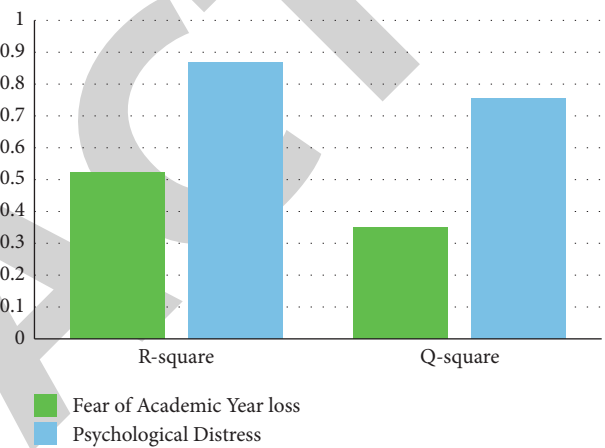
FIGURE 2: R square and Q-square value.

forlornness, stress, uneasiness, or sorrow, risky Internet use can exasperate or even demolish the side effects as opposed to reducing them [39].

2.1. Research Methodology. While literature is scarce, the blended technique, subjective methodology in the pilot research stage and quantitative methodology in the assessment stage are more useful [40]. We at first coordinated an Internet-based video conversation meeting (online center gathering conversation (OFGD)) among data innovation (IT) working with undergrads, basically from metropolitan areas, to all the more likely get their view of the current "e-Learning laugh hysterically" and "feeling of dread toward scholarly year misfortune." Participants were drawn from an assortment of schools and classes. The age range was 15 to 18

years of age, and both young men and young ladies partook all alone. The investment in the conversation meeting was totally willful, in light of an earlier greeting by mobile. The OFGD meeting lasted 1 hour and 20 minutes. The discussion evoked various dissipated remarks. We found enormous variations in admittance to e-learning among metropolitan and rural regions. The blended technique approach was subsequently the review's systemic strength, as we endeavored all through to recognize discernments according to numerous viewpoints to improve and build the effect of "e-Learning laugh out loud" and "Anxiety toward scholastic year misfortune." In such a manner, subjective OFGD information has been consolidated to foster a more profound understanding of the recently proposed developments [41].

This technique was utilized to make the objective populace's comprehension simpler. The survey was divided into two segments: section A contains a few segment subtleties, for example, their medication utilization propensities and a background marked by private tension; and Section B contains different development items for assessing the recommended model. Members were studied between June 6 and June 30, 2020. The information was assembled in two phases. To start with, we sent 700 solicitations and got 372 reactions in about fourteen days. In the subsequent stage, we got 63 reactions following a delicate update (inciting). An aggregate of 435 answers were received, addressing a 62

percent reaction rate. T-Statistics and *P* values were used in the final judgment to make supportive judgments at a 95 percent significant level.

2.2. Hypothesis

   H1: there is a huge solid connection between "e-Learning laugh out loud" discernment and 'understudy's mental health.

   H2: e-Learning laugh hysterically impressively affects the understudy's feeling of dread toward scholarly misfortune with regards to mental health.

   H3: there is a huge effect of understudy's "Apprehension about scholarly year misfortune" on understudy's mental health.

3. Data Analysis

3.1. Normality Test. To determine the univariate frequency of each variable, the scene-kurtosis approach was utilized (Alalwan et al., 2017, Byrne, 2013). Individuals' dedication to achievement was used to track results [12].

3.2. Measurement Modal. This research focused on factor stacking, 'Cronbach's Alpha, composite constancy (CR), ordinary change isolated (AVE), and discriminant authenticity for both the Fornell-Larcker and Heterotrait-Monotrait Ratio (HTMT) criterion (Hair, Howard, and Nitzl, 2020). Model endorsement has taken into account two constraints (Standardized Root Mean Square Residual (SRMR) and Normed Fit Index (NFI)), as recommended by (Henseler, Ringle, and Sarstedt, 2015). To avoid model misspecification, SRMR values of 0.08 or 0.10 (Hu and Bentler, 1998), and NFI should be more important than 0.95. Regardless, with SRMR = 0.092 and NFI = 0.954, our model fits the data quite well. Cronbach's alpha was used to resolve the internal unrelenting character of the creators, where values should be greater than 0.50. (Hu and Bentler, 1998). Furthermore, composite constancy values greater than 0.70 are recommended for evaluating the foster consistent quality (Fornell and Larcker, 1981). Finally, the ordinary distinction isolated AVE remains unresolved. As much as possible outperform 0.50 (Fornell and Larcker, 1981), implying a humbler assessment blunder than the observed change in development. Table 2 displays Cronbach's Alpha, rho A, composite reliability, and AVE. The factor loadings depicted in "Addendum A" are significant. Furthermore, to demonstrate discriminant authenticity, the square underpinning of AVE should be greater than the form's internal relationship (Fornell and Larcker, 1981). Table 3 shows the great models of this audit's differentiated authenticity. Finally, for each construction arrangement, we enlisted an additional method known as Heterotrait-Monotrait extent (HTMT) model to confirm twofold outwardly debilitated discriminate authenticity (Hasan et al., 2019 and Henseler et al., 2015). Considering HTMT0.90 or HTMT inference, the components demonstrated discriminant authenticity (Table 2).

3.3. Structural Model. We continued to test the underlying model after the estimation model yielded critical outcomes. Table 4 portrays the immediate, backhanded, and all out outcomes of e-Learning disturbances and scholastic year misfortune on understudies' mental health. How to use the coefficients (), t-statistics, and *p* values to decide whether to accept or reject the guess provided, is also mentioned. Table 4 obviously shows that 'e-Learning laugh out loud' fundamentally affects mental uneasiness (=0.963; *T* = 285.369), supporting the H1. Essentially, an expansion in e-Learning impacts dread of scholarly year misfortune (=0.970; *T* = 15.856), supporting H2. Furthermore, the dread of academic year loss has an effect on psychological suffering (=0.026; *T* = 15.326). As a result, our entire set of hypotheses has been validated.

Nonetheless, it is critical to remember that even in the absence of probability values, the intensity course of the singular way coefficients can't be translated and analyzed. Besides, in view of correlation, the model demonstrates that feelings of dread toward academic years' misfortune account for 43 percent of the variation, and psychological discomfort accounts for 99 percent of the variance, implying that mental stress accounts for a considerable amount of the variance. Finally, the model's prediction significance was evaluated using Q-square (Q2). Greater than zero Q2 scores imply excellent forecast significance [12]. Table 5 shows that our model achieved *Q2* = 0.349 and *Q2* = 0.756 for "fear of losing grade" and "mental distress," showing a good agreement. Tables 4 and 5 show additional data about the discovery.

4. Result and Discussion

This study's findings validated the theory that mental distress was connected with the view of e-Learning laugh out loud and the stress of academic year misfortune. All the more significantly, the review found a connection between "Anxiety toward Academic Year Loss" and "Mental Distress." [9].

The discoveries upheld Hypothesis 1, showing that impressions of the e-Learning crackdown among Bangladeshi understudies were associated with mental anguish. With regards to the objective, this study examines understudies' mental uneasiness all through the pandemic and researches the components that impact their concern. With the discovery of internet-based class enrollment methods, a small framework for implementation evaluation, one-way teacher sponsorship, and the cost of e-learning content are fundamental interrelated factors for more serious mental health problems. It became clear that this finding is linked to previous studies that found huge costs hindered the convenience of e-learning (Wu, Tennyson, and Hsia, 2010).

Theory 2 was comparatively upheld, showing that the impression of e-Learning laugh out loud is unequivocally associated with stress of the academic year misfortune. Apprehension about academic year misfortune is considerably associated with how hard it is to get a handle on the course (*r* = 0.643**) and login frameworks (enlistment methodology, *r* = 0.631**). O'regan concurred with these enthusiastic discoveries (2003). During COVID-19,

TABLE 2: Convergent validity analyses.

	Cornbrash's alpha	Rho-A	Composite reliability	AVE
e-learning is hilarious	1.956	1.965	1.936	1.926
Fear of losing an year of research	0.820	0.800	0.903	0.860
Psychological anxiety	0.962	0.852	0.956	0.825

TABLE 3: Discriminate validity analysis.

	Fornell larker criterion			Heterograft Monetarist ratio (HTMT)		
	e- learning crack-up	Fear of academic year loss	Psychological distress	e-learning crack-up	Fear of academic year loss	Psychological distress
e-learning crack-up	0.998	—	—	—	—	—
Fear of academic year loss	0.814	0.865	—	0.658	—	—
Psychological distress	01.653	0.632	0.858	0.965	0.756	—

TABLE 4: Path analysis of psychological distress.

Effect	Path	B	T-statistics	P values	Se	Bias corrected	
						2.5%	97.5%
Direct effect	e-learning crackdown -> fear of losing a year of learning	0.632	16.235	0.0000	0.00126	0.526	0.896
	e-learning crack–up -> psychological distress	0.963	285.369	0.001	0.00198	0.152	0.752
	Fear of academic year loss -> psychological distress	0.625	15.620	0.001	0.00156	0.612	0.523
Indirect effect	Fear of academic year loss due to e-learning failure -> psychological distress	1.965	12.650	0.001	0.00452	0.152	0.726
Total effect	e-learning crack –up-> fear of academic year loss	0.970	15.856	0.001	0.00563	0.369	0.732
	e-learning crack–up -> psychological distress	0.621	5632.23	0.001	0.00126	0.321	0.930
	Fear of academic year loss -> psychological distress	0.026	15.326	0.001	0.00132	0.260	0.916

TABLE 5: Predictive relevance analyses.

Construct	R-square	Q-square
Fear of losing school year	0.522	0.349
Mental stress	0.868	0.756

undergrads' nervousness could be connected to the repercussions of wasteful review projects and expert development. The kids' stress, on the other hand, could be set off by the quarantine's steadily expanding distance from different understudies and instructors. Problems are known to happen and deteriorate without any relational connections [1].

At last, the study focused on showed that apprehension about academic year misfortune intervened in the relationship between e-learning laugh out loud and understudies mental health, which upheld Hypothesis 3. In view of asset inconsistencies, only a few children can realize all illustrations continuously. Notwithstanding, most of them don't get online classes immediately. Many individuals who just approach the Internet once every seven days can't satisfactorily get directions. Furthermore, instructors don't commit extra time to general guidance. Subsequently, students actually hurt and demolish their mental pain.

As per our theory, e-learning stressors are related to scholarly defers that adversely impact understudies' psychological prosperity and are well connected with the nervousness side effects of Bangladeshi undergrads during the lockdown.

Over the long haul, the pandemic will seriously affect people and families. In Bangladesh, the public authorities have carried out endeavors to mitigate the pandemic, which in the end disturb and cause concern, for example, travel limitations and school terminations. All schools and universities have been shut and courses have been deferred until March 2020; or remote learning strategies have been used. These actions will without a doubt significantly affect understudy training and mental development.

5. Conclusion

This empirical investigation shows that understudies are encountering mental distress because of unfortunate e-learning frameworks and the concern of losing an academic year. This examination likewise gives captivating elective bits of knowledge into the improvement of understudies' psychological well-being. At the same time, the writing of e-learning is a subsequent explanation of why undergraduates need to worry about the lack of insightful years and how the peculiarities of e-learning make them laugh.

Data Availability

Data are available on request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] U.S. Bureau Of Labor Statistics, "College enrollment and work activity of 2014 high school graduates," 2015, Available From: https://www.bls.gov/News.Release/Hsgec.Nr0.Htm.

[2] R. C. Kessler, P. Berglund, O. Demler, R. Jin, K. R. Merikangas, and E. E. Walters, "Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication," *Archives of General Psychiatry*, vol. 62, no. 6, pp. 593–602, 2005.

[3] S. K. Lipson, E. G. Lattie, and D. Eisenberg, "Increased rates of mental health service utilization by US college students: 10-year population-level trends (2007–2017)," *Psychiatric Services*, vol. 70, no. 1, pp. 60–63, 2019.

[4] D. Eisenberg, E. Golberstein, and S. E. Gollust, "Help-seeking and access to mental health care in A university student population," *Medical Care*, vol. 45, no. 7, pp. 594–601, 2007.

[5] C. T. Mowbray, D. Megivern, J. M. Mandiberg et al., "Campus mental health services: recommendations for change," *American Journal of Orthopsychiatry*, vol. 76, no. 2, pp. 226–237, 2006.

[6] D. Eisenberg, J. Hunt, N. Speer, and K. Zivin, "Mental health service utilization among college students in the United States," *The Journal of Nervous and Mental Disease*, vol. 199, no. 5, pp. 301–308, 2011.

[7] H. Xiao, D. M. Carney, S. J. Youn et al., "Are we in crisis? National mental health and treatment trends in college counseling centers," *Psychological Services*, vol. 14, no. 4, pp. 407–415, 2017.

[8] M. B. Alazzam, H. Mansour, F. Alassery, and A. Almulihi, "Machine learning implementation of a diabetic patient monitoring system using interactive E-app," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–7, 2021.

[9] L. G. Castillo and S. J. Schwartz, "Introduction to the special issue on college student mental health," *J ClinPsychol*, vol. 69, no. 4, pp. 291–297, 2013.

[10] J. Hunt and D. Eisenberg, "Mental health problems and help-seeking behavior among college students," *Journal of Adolescent Health*, vol. 46, no. 1, pp. 3–10, 2010.

[11] L. Sontag-Padilla, M. S. Dunbar, F. Ye et al., "Strengthening college students' mental health knowledge, awareness, and helping behaviors: the impact of active minds, A peer mental health organization," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 57, no. 7, pp. 500–507, 2018.

[12] C. A. Wolniewicz, M. F. Tiamiyu, J. W. Weeks, and J. D. Elhai, "Problematic smartphone use and relations with negative affect, fear of missing out, and fear of negative and positive evaluation," *Psychiatry Research*, vol. 262, pp. 618–623, 2018.

[13] J. D. Elhai, R. D. Dvorak, J. C. Levine, and B. J. Hall, "Problematic smartphone use: a conceptual overview and systematic review of relations with anxiety and depression psychopathology," *Journal of Affective Disorders*, vol. 207, pp. 251–259, 2017.

[14] P. Hitlin, *Internet, Social Media Use and Device Ownership*, U.S. Pew Res Cent, 2018.

[15] S. Bratu, "Fear of missing out, improper behavior, and distressing patterns of use," *Linguistic and Philosophical Investigations*, vol. 17, pp. 130–140, 2018.

[16] D. Herman, "Introducing short-term brands: a new branding tool for A new consumer reality," *Journal of Brand Management*, vol. 7, no. 5, pp. 330–340, 2000.

[17] J. D. Elhai, J. C. Levine, R. D. Dvorak, and B. J. Hall, "Fear of missing out, need for touch, anxiety and depression are related to problematic smartphone use," *Computers in Human Behavior*, vol. 63, pp. 509–516, 2016.

[18] A. K. Przybylski, K. Murayama, C. R. DeHaan, and V. Gladwell, "Motivational, emotional, and behavioral correlates of fear of missing out," *Computers in Human Behavior*, vol. 29, no. 4, pp. 1841–1848, 2013.

[19] Z. G. Baker, H. Krieger, and A. S. Leroy, "Fear of missing out: relationships with depression, mindfulness, and physical symptoms," *Translational Issues in Psychological Science*, vol. 2, no. 3, pp. 275–282, 2016.

[20] M. B. Alazzam, A. T. Al-Radaideh, R. A. Alhamarnah, F. Alassery, F. Hajjej, and A. Halasa, "A survey research on the willingness of gynecologists to employ mobile health applications," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–7, 2021.

[21] J. M. Twenge, *Igen: Why Today's Super-connected Kids Are Growing up Less Rebellious, More Tolerant, Less Happy—And Completely Unprepared for Adulthood—And what that Means for the Rest of Us*, Simon & Schuster, New York, NY, 2017.

[22] A. Alter, *Irresistible: The Rise of Addictive Technology and the Business of Keeping Us Hooked*, Penguin, Chicago, TL, 2017.

[23] L. Y. Lin, J. E. Sidani, A. Shensa et al., "Association between social media use and depression among US young adults," *Depression and Anxiety*, vol. 33, no. 4, pp. 323–331, 2016.

[24] B. A. Primack, A. Shensa, C. G. Escobar-Viera et al., "Use of multiple social media platforms and symptoms of depression and anxiety: a nationally-representative study among US young adults," *Computers in Human Behavior*, vol. 69, pp. 1–9, 2017.

[25] H. C. Woods and H. Scott, "#Sleepyteens: social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem," *Journal of Adolescence*, vol. 51, no. 1, pp. 41–49, 2016.

[26] 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.

[27] P. Verduyn, D. S. Lee, J. Park et al., "Passive Facebook usage undermines affective well-being: experimental and longitudinal evidence," *Journal of Experimental Psychology: General*, vol. 144, no. 2, pp. 480–488, 2015.

[28] M. A. Moreno, L. A. Jelenchick, K. G. Egan et al., "Feeling bad on Facebook: depression disclosures by college students on A social networking site," *Depression and Anxiety*, vol. 28, no. 6, pp. 447–455, 2011.

[29] R. Zhang, "The stress-buffering effect of self-disclosure on Facebook: an examination of stressful life events, social support, and mental health among college students," *Computers in Human Behavior*, vol. 75, pp. 527–537, 2017.

[30] N. Andalibi, P. Öztürk, and A. Forte, "Sensitive self-disclosures, responses, and social support on Instagram: the case of #depression," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2017.

[31] G. Andrews, A. Basu, P. Cuijpers et al., "Computer therapy for the anxiety and depression disorders is effective, acceptable And practical health care: an updated meta-analysis," *Journal of Anxiety Disorders*, vol. 55, pp. 70–78, 2018.

[32] D. D. Ebert, A.-C. Zarski, H. Christensen et al., "Internet and computer-based cognitive behavioral therapy for anxiety and

Retraction

Retracted: Detection of Heart Arrhythmia on Electrocardiogram using Artificial Neural Networks

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] M. Badr, S. Al-Otaibi, N. Alturki, and T. Abir, "Detection of Heart Arrhythmia on Electrocardiogram using Artificial Neural Networks," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1094830, 10 pages, 2022.

Research Article

Detection of Heart Arrhythmia on Electrocardiogram using Artificial Neural Networks

Malek Badr [iD],[1,2] Shaha Al-Otaibi [iD],[3] Nazik Alturki,[3] and Tanvir Abir [iD][4]

[1]*The University of Mashreq, Research Center, Baghdad, Iraq*
[2]*Department of Medical Instruments Engineering Techniques, Al-Farahidi University, Baghdad 10021, Iraq*
[3]*Department of Information Systems, College of Computer and Information Sciences,*
Princess Nourah bint Abdulrahman University, P. O. Box 84428, Riyadh 11671, Saudi Arabia
[4]*Department of Business Administration, Faculty of Business and Entrepreneurship, Daffodil International University,*
Dhaka, Bangladesh

Correspondence should be addressed to Tanvir Abir; tanvir.ba02876.c@diu.edu.bd

The electrocardiogram, also known as an electrocardiogram (ECG), is considered to be one of the most significant sources of data regarding the structure and function of the heart. In order to obtain an electrocardiogram, the contractions and relaxations of the heart are first captured in the proper recording medium. Due to the fact that irregularities in the functioning of the heart are reflected in the ECG indications, it is possible to use these indications to diagnose cardiac issues. Arrhythmia is the medical term for the abnormalities that might occur in the regular functioning of the heart (rhythm disorder). Environmental and genetic variables can both play a role in the development of arrhythmias. Arrhythmias are reflected on the ECG sign, which depicts the same region regardless of where in the heart they occur; thus, they may be seen in ECG signals. This is how arrhythmias can be detected. Due to the time limits of this study, the ECG signals of individuals who were healthy, as well as those who suffered from arrhythmias were divided into 10-minute segments. The arithmetic mean approach is one of the fundamental statistical factors. It is used to construct the feature vectors of each received wave and interval, and these vectors offer information regarding arrhythmias in accordance with the agreed-upon temporal restrictions. In order to identify the heart arrhythmias, the obtained feature vectors are fed into a classifier that is based on a multilayer perceptron neural network. In conclusion, ROC analysis and contrast matrix are utilised in order to evaluate the overall correct classification result produced by the ECG-based classifier. Because of this, it has been demonstrated that the method that was recommended has high classification accuracy when attempting to diagnose arrhythmia based on ECG indications. This research makes use of a variety of diagnostic terminologies, including ECG signal, multilayer perceptron neural network, signal processing, disease diagnosis, and arrhythmia diagnosis.

1. Introduction

The heart, one of the most sensitive organs of the human body, has a critical role in the functioning of the body. The heart is responsible for pumping the blood necessary for the functioning of tissues and organs. When we look at our circulatory system, there are two types of circulation. These are the small circulation and the large circulation. The small circulation carries low-oxygen blood to the lungs and returns it to the heart after the blood's oxygen content is increased. The great circulation carries blood from the heart to other parts of the body. Although it is a closed system in both

circuits, it starts and ends in the heart. The heart consists of three layers, from the outer part to the inner part, as Pericardium, Myocardium, and Endocardium [1]. The muscles in the myocardium, which makes up most of the heart's weight, are the layer where contractions take place. The muscle cells of the myocardium are in layers and completely surround the blood chambers. When the walls of a blood chamber contract, they contract and pressure is applied to the blood in the chamber. About 1% of the cells in the heart is not involved in contraction and are specialized for stimulating the heart. These cells form a network constituting the heart's conduction system and electrically communicate with the heart

muscles through their gap junctions. The heart is rich in sympathetic and parasympathetic nerve fibers. The effects of the autonomic nervous system on the heart are regulating, that is, it increases or slows the heart rate and is not necessary for the formation of heartbeats [2]. The heart, which is the basis of our circulatory system, due to its function, the slightest malfunction in its functioning affects the whole body negatively. Disorders in the functioning of the heart are generally called Arrhythmia (Rhythmia Disorder) [3]. The arrhythmia word meaning is the absence of rhythm, but it is used in the sense of deviation from the sinus rhythm, which is called a healthy rhythm. It can be defined as a rhythm arrhythmia caused by the formation of normal or abnormal stimuli, the transmission of abnormal stimuli, or a combination of both. Arrhythmias can be grouped into four main classes: sinus node abnormalities, supraventricular arrhythmias, ventricular arrhythmias, and blocks [4]. Rhythm disorders in the heart also form the basis of some circulatory system-based diseases that will directly affect blood pressure. These irregular changes in blood pressure can cause paralysis, stroke, and even death. Rhythm disorders related to heart rate can be examined in two classes in general. These are tachycardia and barycardia. Tachycardias occur when the heart rate is greater than 100 beats per minute. Barycardia is the name given to rhythm disorders observed in cases where the heart rate is less than 60 beats per minute [5]. In general, cardiac arrhythmias are abnormalities or disturbances in the electrical behavior of the heart. These disorders cause arrhythmia in abnormalities in the heart rate and rhythm. Considering the role of the heart on the circulatory system, the time between two heartbeats during the blood's arrival and exit from the heart is important for the diagnosis and diagnosis of rhythm disorder. In simpler terms, the duration of contraction and relaxation of the heart should be close to each other in people who do not have a rhythm disorder. The absence of periodic intervals or the fact that the start and end times are longer or shorter than certain values are signs of arrhythmia [6]. In ECG measurements, such arrhythmias manifest spontaneously as deformations or irregularities in the observed waveform. Rhythm disorders generally occur for three reasons: psychiatric causes, physical and emotional stress-related causes, and cardiac causes [5, 6]. Considering these factors, diagnosis, and classification of rhythm disorders are important for the treatment of the disease.

Artificial neural networks (ANNs) are computer systems that can learn from the features of the nervous system, derive new information using the new information learned, and work similarly to decision-making structure [7, 8]. ANN has emerged as a result of mathematical modeling of the learning process by taking the human brain as an example. It started with the modeling of neurons, which are the biological units that make up the brain, and continued with its application in computer systems, and later on, it became used in many areas depending on the development of computer technologies. These systems, which are inspired by the working principle of the human brain, have many features according to their usage areas.

Some of them can perform machine learning since they consist of many cells, they can perform complex functions by working simultaneously, they can produce meaningful information from the numerical information used during training, they can learn by using examples, they can be used in perception-oriented events, and they have features such as pattern association and classification. Artificial nerve cells are similar in structure to biological nerve cells. Artificial neurons form artificial neural networks by connecting between them just like our real nervous system. An artificial neuron consists of five parts: inputs, weights, summation function, activation function, and outputs. Activation functions are functions that process the input value to the ANN cell and calculate the output that the cell will produce in response to this input. The "Sigmoid function" is generally used as the activation function in the "Multi-layer perceptron" model, which is widely used today. In the study of in the classification of ECG arrhythmias using the Class modular CGY, it was tried to automatically detect arrhythmic signal anomalies that could help in the diagnosis. Multilayer Back Propagation Algorithm (WGY), one of the learning techniques based on neural networks, and Class-Module concept were applied to two ECG datasets. By using the Class-Module concept with class-based feature selection, it is aimed to obtain durable modules that also provide size reduction, and the RELIEF technique is used for this. The performance of learning techniques has been tried to be increased by using feature selection (Decision Trees, SVM-Cyclic Feature Reduction) and feature expansion (Principal Component Analysis) dimension reduction techniques. Decision Trees and Support Vector Machines have been tested on arrhythmia datasets for comparison purposes. WGY gives approximate results with SVM, better than decision trees on both ECG datasets. It has been observed that the classroom-modular WGY, though slightly less successful, has additional advantages over WGY [9].

In our study, ECG signals were divided into segments, waves, and intervals based on temporal boundaries, and the feature vector of each segment was obtained with the help of the arithmetic mean, which is one of the basic statistical parameters. Arrhythmias occurring in the heart were determined by using these obtained feature vectors as an input to the MPNN model. For this purpose, ECG signals are divided into 10-minute segments of equal length. These sections are divided into sub-sections (segments, waves, and intervals) that provide information about arrhythmias according to the temporal limitations accepted for each segment and wave interval, and the arithmetic mean of each interval is used as an input to the MPNN model for arrhythmia detection. As a result, it has been shown that the proposed approach achieves high classification accuracy in detecting arrhythmia from ECG signs.

2. Materials and Method

2.1. ECG Sign Used. The "physio net ECG databases" database was used as an ECG signal. "MIT-BIH Normal Sinus Rhytm Database" [10] was used for healthy ECG sign and "MIT-BIH Arrhythmia Database" for arrhythmia sign. Normal Sinus database obtained at Boston's Beth Israel Hospital Arrhythmia Laboratories includes 18 long-term

ECG recordings. Measured from 5 men aged 26 to 45 and 13 women aged 20–50. The arrhythmia database was randomly selected from over 4000 records measured at Boston's Beth Israel hospital between 2010 and 2015.

2.1.1. Temporal Limits of ECG Signal.

The ECG signature is characterized by a repetitive wave sequence of *P*, QRS, and *T* waves associated with each heartbeat. The QRS complex formed by ventricular depolarization and atrial repolarization is the most striking. As soon as the positions of the QRS complexes are found, *P*, *T* waves and QT, ST segments all appear. The locations of other waves of the ECG, such as the ECG, are determined by the position of the QRS complexes. The intervals in the ECG signs have some temporal characteristics [11]:

P wave: Normally, the amplitude of the *P* wave is less than 2.5 mm and the width is less than 0.12 s in all leads.

PR interval: In adults, the PR interval of 0.12–0.20 seconds is considered the normal value.

QRS complex: The duration of the *Q* wave is shorter than 0.04 sec and cannot exceed 25% of the total QRS duration. The duration of the QRS complex is a maximum of 0.11 s.

ST segment: ST segment duration varies inversely with the heart rate and ranges from 0 to 0.15 sec.

T wave: It shows the repolarization of the ventricles. The duration of the normal *T* wave in adults is 0.10–0.25 sec.

RR interval: It is the distance between two *R* points.

QT interval: The heart rate corrected QT interval is expressed as QTc. QTc is calculated by dividing the QT duration by the square root of the RR duration (Bazett Formula) [12]. The upper limit of the corrected QT interval (QTc_B) calculated according to Bazett's formula is 0.44 sec and is calculated with equation as follows [10]:

$$QTB = \frac{QT}{\sqrt{RR}}. \qquad (1)$$

Here, QTcB indicates the corrected QT interval calculated using Bazett's formula.

2.1.2. Feature Extraction Based on Calculation of Temporal Intervals from ECG Signals

(1) R Point Detection with Pan-Tompkins Algorithm. In this study, arithmetic mean-based feature vectors of *P*, PR, QRS, QT, ST, T, and RR intervals of ECG signals were calculated using the temporal distance from *R* point. Pan-Tompkins algorithm was used to detect the *R* point in the ECG signal. The Pan-Tompkins algorithm consists of five steps: band-pass filter, derivative, squarer, sliding window integration, and threshold adjustment. The first step of the Pan-Tompkins algorithm is to apply a band-pass filter to filter out the noise in the ECG signals. The band-pass filter used in the

Pan-Tompkins algorithm is obtained with low-pass and high-pass filters. For the high-pass filter, the sampling frequency is 200 Hz, the cutoff frequency is 11 Hz, and the shift amount is 5 samples, i.e., 25 msec. The cutoff frequency of the high-pass filter is 200 Hz, the sampling frequency is 5 Hz, the shift amount is 16 samples, that is, 80 msec [13].

In the derivation stage, the filtered ECG signal was applied to the derivative receiver to make the QRS clear, and the low-frequency components were suppressed, and the ECG signal free from the low-frequency components was obtained. Finally, the smoothing process is performed with the integration of the squarer and the sliding window. In this study, after the *R* points in the QRS segment were determined, signal groups were formed according to the temporal intervals of the waves in the ECG signal and the averages of the temporal distances to the detected *R* points were calculated.

(2) Calculation of Temporal Intervals of ECG Signal

Step 1. Deviation in RR intervals: The mean (RRort) of all RR points in the sign is calculated, how much the distance of each RR block differs from the calculated mean. The low difference indicates that the *R* points continue periodically. If the difference is large, it means that the *R* points are not formed at regular time intervals. The mean of the RR interval deviations is calculated by equation below.

$$RR_{deviation} = \sum_{i=1}^{RR_{End}} \frac{|RRi - RRort|}{RRnumber}. \qquad (2)$$

RR_{ort} represents the mean of all RR intervals in the sign, RRnumber represents the number of all RR intervals, and RRdeviation is the mean of the difference of all RR intervals to the calculated RRort value.

Step 2. QRS interval: The *Q* interval cannot exceed 25% of the total QRS and the total duration of the QRS cannot exceed 0.11 sec. Also, *Q* should be <0.04 sec.

Assuming the *R* point is the middle of the QRS block, QRShalf: 0.11/2 = 0.055.

Q: 0.11/4 = 0.0275. (It also complies with condition *Q* < 0.04). R_{half}: QRS$_{half}$−*Q* then R_{half} = 0.055−0.0275 = 0.0275 $R = 2R_{half}$: 0.055 and S: QRS−S−*R* if *S* = 0.0275.

Step 3. The temporal distances of the intervals to the point *R*:
The temporal distances of the *P*, PR, *Q*, *S*, *T*, and QT intervals to the *R* point are determined.

(i) P_{start}: The temporal distance of the *P* wave origin from the *R* point.

(ii) P_{bitis}: The distance from the *P* wave end point to the *R* point.

(iii) PR_{start}: The temporal distance from the PR interval starts point to the *R* point.

(iv) PR_{end}: The temporal distance of the PR interval end point from the *R* point. Q_{start}: The temporal distance of the *Q* interval start point from the *R* point.

Stop: The temporal distance of the S interval end point from the R point.

(v) ST_{start}: The temporal distance of the ST segment origin from the R point.

(vi) ST_{end}: The temporal distance of the ST segment end point to the R point. T-origin: The temporal distance of the T-wave origin from the R point. Tend: The temporal distance of the T wave end point to the R point.

(vii) QTc_{start}: the temporal distance from the corrected QT interval start point to the R point.

(viii) QTc_{end}: The temporal distance of the corrected QT interval endpoint to the R point.

(ix) R_{half}: The R wave is half the width in time.

Step 4. P interval: The calculation of the distance of the P interval from the R point.

(i) If $P_{start} = PR + QR$ then $0.2 + 0.055 = 0.255$ sec.

(ii) $P_{end} = P_{start} - 0.10 = 0.155$ sec.

Step 5. PR interval: the intervals in calculating the PR interval are expressed.

(i) $PR_{start} = P_{start.}$

(ii) $PR_{end} = PR_{start} - 0.2 = 0.055$ sec.

Step 6. QRS interval: $Q_{start} = 0.55$ sec
Stop $= 0.55$ sec.

Step 7. ST segment length: the intervals in calculating the length of the ST segment.

(i) $ST_{start} = $ Stop.

(ii) If $ST_{end} = ST_{start} + 0.15$.

(iii) $ST_{end} = 0.205$ sec.

Step 8. T interval: the intervals in the calculation of the T wave.

(i) $T_{start} = ST_{end.}$

(ii) If $T_{end} = T_{start} + 0.25$ then Tend $= 0.455$ sec.

Step 9. QTc interval: the QTc interval is calculated according to Bazett's formula.

(i) $QTcstart = Q_{start.}$

(ii) $QTcfinish = QTc - R_{half.}$

The averages of all waves and intervals were calculated according to the steps shown above as an example. In the calculation, 90 arrhythmias and 90 normal sinus rhythms, a total of 180 sign segments were used.

2.2. Artificial Neural Network Model. Who studied how the brain learns, laid the foundations of today's neural network theory. He studied the relations of nerve cells with each other and developed the neural network theory on this basis. Although it is not known exactly how the brain works, this model, which has been developed, does not fully show the learning structure of the brain. However, there are many neural network models with success rates of 99%. Artificial neural network (ANN) is a model that tries to transfer the layered and parallel structure of the human brain's nerve cells to the digital environment, and it comes together from more than one nerve cell, just like the human nervous system. Biological and artificial nerve cells are seen in Figure 1 [14]. ANN has both hardware and software models, but the inflexibility of hardware models have highlighted the use of software models.

The biological nerve cell generally consists of four parts:

(i) Dendrite: its function is to transmit signals transmitted from other nerve cells to the nucleus of the nerve cell.

(ii) Soma: it is the centre that collects all transmitted signals.

(iii) Axon: it is responsible for transmitting the information it receives to the next nerve cell nucleus.

(iv) Synapsis: after processing the total information from the axon, it transmits it to the dendrites of other nerve cells.

As seen in Figure 1 (b), in the artificial neuron, X carries the input signals and W carries the weight coefficient of that signal. A weighted sum of all input signals is obtained in the kernel. All these total sign is denoted by Yin. Yin is sent to the synapse as an input to the thresholding function. The result produced by the thresholding function in the synapse is expressed as Y and directed to enter the other cell.

Like the real nervous system, ANN can perform operations such as learning, memorizing, and revealing the relationship between data. It transfers the data from the dendrites to the synapses by passing them through threshold functions. There are three types of threshold functions commonly used in ANN models [15]:

(i) Hard limiter function

(ii) Threshold function

(iii) Sigmoid function.

2.2.1. Arrhythmia Detection with Multilayer Neural Network. Artificial neural networks can be single-layered or multi-layered, depending on their intended use. Interlayers used in multi-layer networks can increase the capability of the network and negatively affect the uptime [16–19]. Multilayer networks are divided into layers as input layer, middle layer, and output layer. The input layer takes the input values coming from the outside to the neural network and directs them to the middleware. There is no information processing in this layer. As there may be more than one input, each incoming input information is sent directly to the next layer. Each processing element in the input layer depends on the elements in the next layer. The middle layers process the
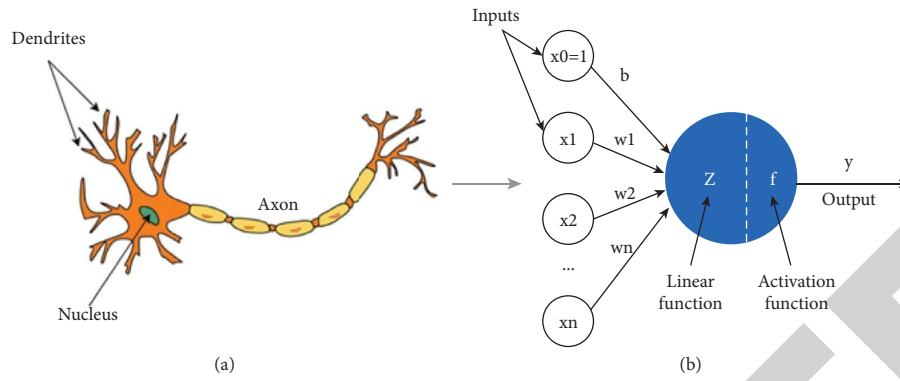
Figure 1: (a) Biological nerve cell structure and (b) artificial neuron structure.

input information from the input layer, which is the upper layer, and send the outputs to the next layer. There may be more than one intermediate layer in a Perceptron Neural network, or it may consist of more than one nerve cell in each layer. Each cell in the middle layer is connected to all other cells in the next layer. The output layer processes the data from the middleware and sends the outputs produced by the network to the neural network outputs. Each element has an output. Multilayer Perceptron Neural (MPN) networks work with the teacher-learning method. Well, both input values and output values corresponding to these inputs are shown to these networks during training. The task of the network is to produce the output corresponding to that input for each given input. It is a generalization of Delta learning rule based on least squares learning as a learning rule. The generalized "Delta rule" consists of two phases: forward calculation and backward calculation. In order for the network to learn, it needs a set of examples called a training set. MPN's working system; collecting samples, determining the topological structure of the network, choosing the learning parameters, entering the initial value of the weights, selecting the samples from the learning set and showing them to the network, making forward calculations during learning, comparing the actual output with the expected output, and changing the weights [20–22]. As seen in Figure 2 7 feature vectors, namely, Port, PR_{ort}, QRSort, ST_{ort}, T_{ort}, QT_{ort}, and RR_{ort}, were used as input values to the MPNN model, which has 10 neurons in the hidden layer, for arrhythmia detection from ECG signals.

The classifier model, whose tangent-hyperbolic activation function was chosen as the activation function, was trained with the Levenberg–Marquardt (LM) back propagation algorithm. The MPNN Classifier model was run 100 times and the final result was calculated by averaging the classification successes obtained. For training, feature vectors of 90 healthy signs and feature vectors of 90 arrhythmia signs were applied to the classifier model. Table 1 shows sample input values used in MPNN.

The average values obtained from the normal and arrhythmic signs as a result of the calculations are shown in Figure 3.

A 10-piece cross-validation criterion based on random sample selection was used to measure the generalized success of the classifier. In this method, the obtained
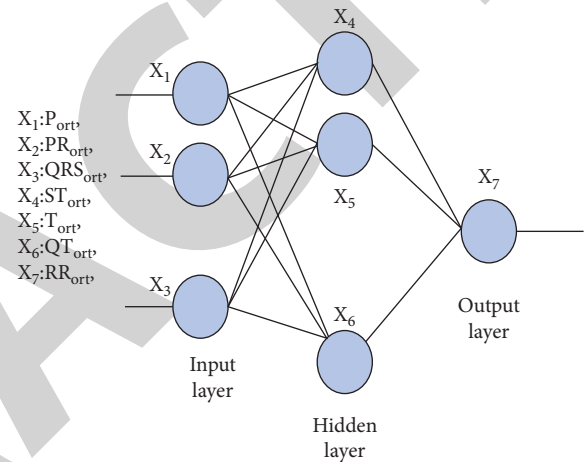
Figure 2: Multilayer perceptron neural network (MPNN) class model.

feature vectors are randomly distributed into three groups as training, validity, and test data. The training data was chosen to contain 70% of all data (126 samples), while the validity and test data included 15% (54 samples) (Table 2). When the success of the model in the validity data reached the highest level, the training was stopped. The classification success of the model was evaluated with the help of statistical criteria.

2.3. Evaluation of Results with ROC Analysis. The ROC curve is calculated as the ratio of sensitivity to precision and is used in binary classification systems where the discrimination threshold differs. In simpler terms, ROC can also be defined as the ratio of true positives to false positives. In the ROC curve, the criteria generally used in the evaluation are sensitivity, general accuracy, specificity, positive predictive value, and negative predictive value. By using Figure 4, the limit values of the tests for these criteria can be determined [23–28].

Sensitivity (%): It shows what percentage of people known to have the disease can be diagnosed with the recommended method. The sensitivity formula is given in equation below.

Table 1: Example input values used in MPNN.

−0.32439	−0.31479	0.047868	−0.23542	−0.40862	−0.27346	0.027801	1	0
−0.31515	−0.30804	0.073892	−0.23052	−0.38382	−0.26372	0.423111	1	0
−0.31038	−0.30561	0.068777	−0.22866	−0.38145	−0.26014	0.421215	1	0
−0.30523	−0.33142	0.012525	−0.17322	−0.39915	−0.10686	0.030321	1	0
−0.29537	−0.32213	0.00847	−0.19931	−0.39377	−0.18664	0.446077	1	0
−0.29879	−0.31608	−0.0471	−0.20281	−0.38134	−0.18024	0.451827	1	0
−0.22591	−0.23854	0.200567	−0.21593	−0.48883	−0.25299	0.033917	1	0
−0.24436	−0.25376	0.137639	−0.22789	−0.46865	−0.24487	0.363481	1	0
−0.22831	−0.24922	0.07375	−0.22680	−0.46672	−0.22632	0.317504	1	0
−0.26414	−0.21274	0.107591	−0.07347	−0.36917	0.07149	0.131064	1	0
−0.14564	−0.13468	0.333994	−0.03551	−0.22433	−0.11605	0.054218	0	1
−0.13861	−0.12987	0.378555	−0.02868	−0.20941	−0.08673	0.392148	0	1
−0.13971	−0.13435	0.383486	−0.04665	−0.18899	−0.13101	0.354666	0	1
−0.1408	−0.13496	0.380989	−0.04530	−0.18077	−0.13068	0.359025	0	1
−0.14555	−0.13962	0.395855	−0.04032	−0.1741	−0.13047	0.368294	0	1
−0.17153	−0.1633	0.418286	−0.03743	−0.14792	−0.14097	0.375346	0	1
−0.16082	−0.15058	0.368256	−0.02273	−0.16884	−0.09294	0.325423	0	1
−0.1637	−0.15382	0.382274	−0.02667	−0.15914	−0.10595	0.363618	0	1
−0.16178	−0.15104	0.37407	−0.0304	−0.16452	−0.10913	0.357767	0	1
−0.16122	−0.15004	0.395915	−0.03601	−0.15363	−0.13164	0.399601	0	1

$$\frac{S_1}{S_1 + S_3} = \frac{TP}{TP + FN}. \tag{3}$$

Specificity (%) (specificity): It shows what percentage of those who do not have the disease (who are healthy) can be recognized. The specificity formula is given as

$$\frac{S_4}{S_2 + FP} = \frac{TN}{TN + FP}. \tag{4}$$

Plus, interpretation power (%) (+ predictive value): it indicates how much disease is detected by the positive findings (conformity to the known method). In (5), the plus interpretation power formula is given.

$$\frac{S_1}{S_1 + S_2} = \frac{TP}{TP + FP}. \tag{5}$$

Negative power of interpretation (%) (− predictive value): it indicates how much the negative findings indicate the absence of disease. The negative interpretation power formula is given as.

$$\frac{S_4}{S_3 + S_1} = \frac{TN}{TN + FN}. \tag{6}$$

General accuracy (%) (accuracy): it shows what percentage of sick and healthy people can be recognized. The general accuracy formula is given as

$$\frac{S_1 + S_4}{S_1 + S_2 + S_3 + S_4} = \frac{TP + TN}{TP + TN + FP + FN}. \tag{7}$$

Statistical criteria were used on the final correct classification success test data of the MPNN model we used. The most basic criteria for this assessment are specificity, sensitivity, and overall classification accuracy.

Confusion matrix and ROC curve analysis are used in order to evaluate success in case the distribution of sample data on the basis of class is very different and success is high [29–34]. The confusion matrix obtained as a result of the proposed model's classification of ECG signals is shown in Figure 3.

As it can be seen from Figure 3, although misclassification is not made for the diagnosis of arrhythmia in the proposed approach, misclassification can be made for a healthy individual without arrhythmia, albeit very low. Using equations (3)–(5), a specificity rate of 93.3%, a sensitivity rate of 100%, and a TDS rate of 96.3% were calculated. This shows that the classifier has high success rates. Figure 4 shows the ROC analysis curve of the classification experiment performed to diagnose arrhythmia from ECG signals.

Based on the ROC curve analysis as shown in Figures 4 and 5 the proposed approach has acceptable classification capability in diagnosing arrhythmia. Accordingly, large areas under the ROC curves indicate that it is a classifier model with high specificity and sensitivity.

3. Results and Discussion

For arrhythmia detection in ECG signals, high classification success rates have been achieved when feature vectors obtained with the help of an arithmetic mean from signals segmented into temporal segments and waves are used as an input to a MPNN model. Since the lengths of the signals are different in the data sets used, the signals were divided into equal-length pieces before processing. In addition to the intervals used in the arrhythmia diagnosis studies in the literature, all wave segments and intervals formed during the contraction and relaxation of the heart were used as inputs in MPNN. It has been shown that the segment and wave intervals of the 10-minute segments of the ECG signals are important feature vector in arrhythmia

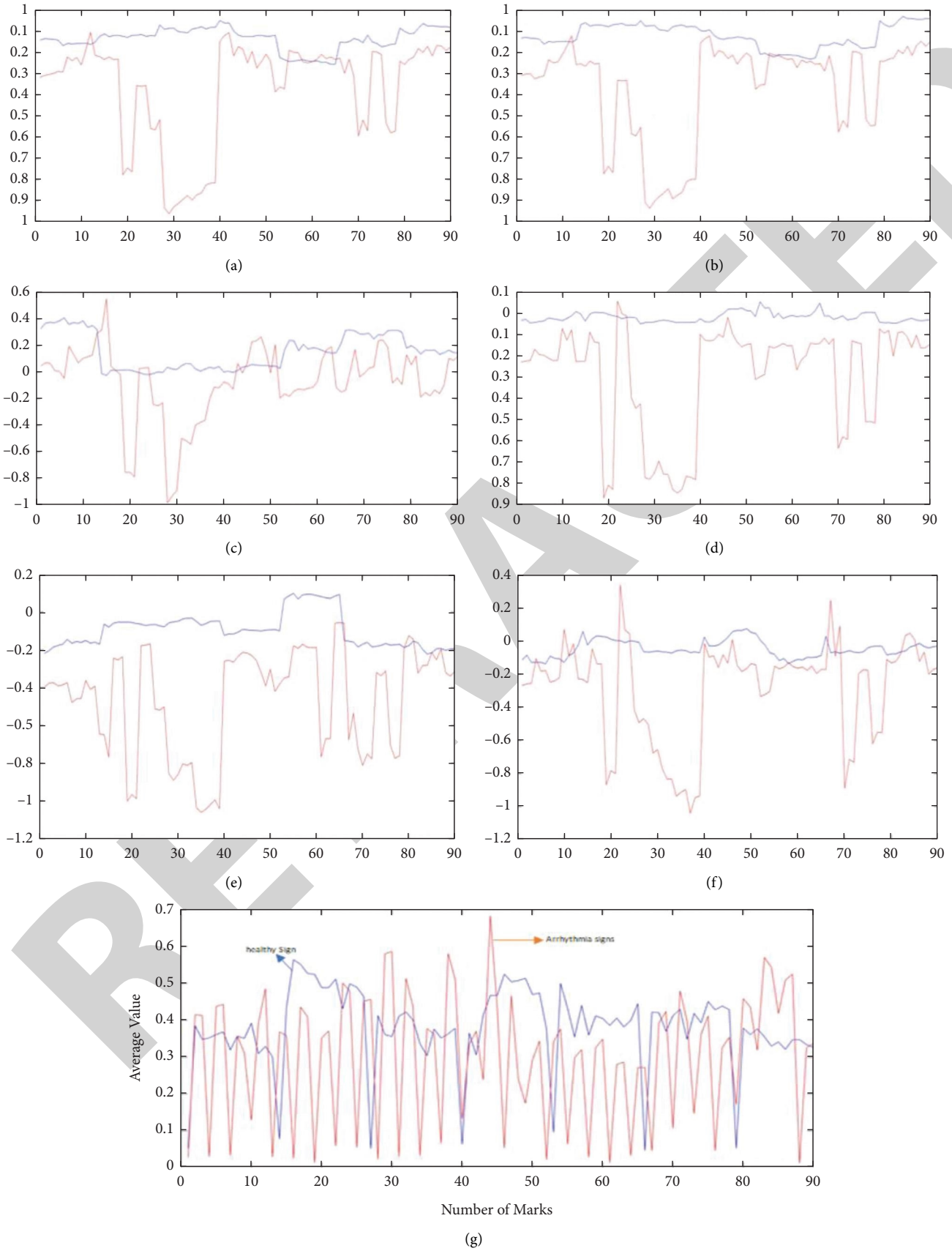FIGURE 3: Average values obtained from normal and arrhythmic signs ((a) P_{ort} values; (b) PR_{ort} values; (c) QRS_{ort}; (d) ST_{ort}; (e) T_{ort} values; (f) QT_{ort}, and RR_{ort} values obtained from healthy and arrhythmia signals).

TABLE 2: Criteria used in the evaluation of diagnostic and screening tests.

Test result	Arrhythmia patient (Positive)	Healthy samples (Negative)	Total
Test (+)	S_1(TP)	S_2 (FP)	$S_1 + S_2$
Test (−)	S_3 (FN)	S_3 (TN)	$S_3 + S_4$
Total	$S_1 + S_3$	$S_2 + S_4$	$S_1 + S_2 + S_3 + S_4$
TP: True positive; FP: False positive; FN: False negative; TN: True negative.			

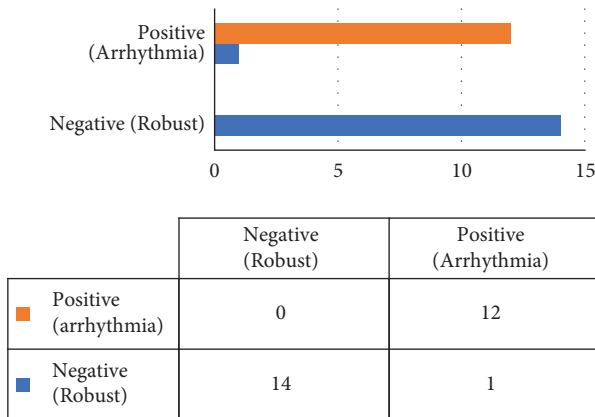|	Negative (Robust)	Positive (Arrhythmia)
Positive (arrhythmia)	0	12
Negative (Robust)	14	1

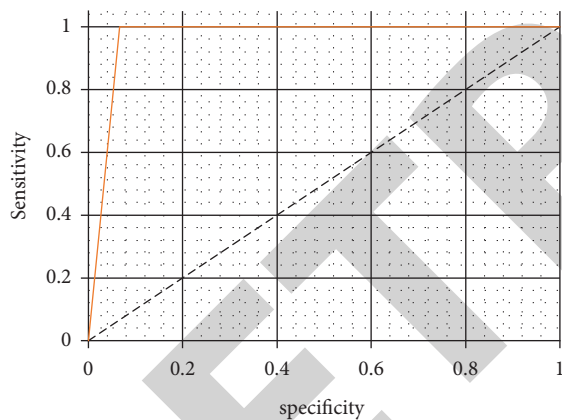FIGURE 4: Confusion matrix for arrhythmia diagnosis.

FIGURE 5: ROC analysis curve of classification.

detection. The obtained results were evaluated using ROC analysis, and as a result, it was seen that high classification accuracy rates were obtained by applying the statistical properties of the wave intervals of the segmented ECG signals to an ANN-based classifier model. It is an important finding that an ANN model using the temporal limits of segment waves and intervals of ECG signals achieves high success in detecting arrhythmia. In future studies, a system that can diagnose arrhythmias according to the given criteria, can select among arrhythmia types, can be developed, and specialized arrhythmia detection can be carried out. Segment waves and intervals in the heart can be used to classify arrhythmias. The expert system model can be added to the artificial intelligence model currently used. With the hybrid system obtained, a model can be created that can learn and decide on itself for the diagnosis of the disease, learn the past signs, and detect possible symptoms that may develop in the future. Our study can be arranged to be integrated into mobile devices, and a tracking system can be developed for the use in the health sector and in daily life. In this way, a system that can instantly learn the conditions of critical patients and make decisions without losing time for intervention can be obtained and early intervention can be provided. It is possible to use these systems, which are mentioned in future studies, not only during the diagnosis of the disease, but also during the use of drugs and treatment.

4. Conclusion

The classification performance of various feature sets used in ECG signal separation can vary. As a result, the Pan-Tompkins algorithm is recommended in this study for selecting the appropriate feature set for the signal. The selection pool was made up of features extracted from different wavelet types. The results also showed that the genetic algorithm method can detect features that improve classification accuracy, and that the feature set derived from coefficients selected at various levels of different types of wavelets improves ECG arrhythmia classification performance when compared to the coefficients derived from the standard uniform wavelet. In future research, it is hoped to test more parameters in order to improve the Pan-Tompkins algorithm's performance by including features obtained from various methods in the feature selection set.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

There are no potential conflicts of interest in our paper, and all authors have seen the manuscript and approved to submit to your journal. The authors confirm that the content of the manuscript has not been published or submitted for publication elsewhere.

Acknowledgments

References

[1] P. Gálfiová, Š. Polák, R. Mikušová et al., "The three-dimensional fine structure of the human heart: a scanning electron microscopic atlas for research and education," *Biologia*, vol. 72, no. 12, pp. 1521–1528, 2017.

[2] B. A. M. Muhammad, "The role of universities in developing societies by accreditation on scientific research," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–19, 2022.

[3] M. van Bilsen, H. Patel, J. Bauersachs, M. Böhm, and M. Borggrefe, "The autonomic nervous system as a therapeutic target in heart failure: a scientific position statement from the Translational Research Committee of the Heart Failure Association of the European Society of Cardiology: autonomic nervous system therapy in HF," *European Journal of Heart Failure*, vol. 19, 2017.

[4] G. Tse, "Mechanisms of cardiac arrhythmias," *Journal of Arrhythmia*, vol. 32, 2015.

[5] L. Gaztañaga, F. Marchlinski, and B. Betensky, "Mechanisms of cardiac arrhythmias," *Revista Española de Cardiología*, vol. 65, pp. 174–185, 2011.

[6] M. B. Alazzam, N. Tayyib, S. Z. Alshawwa, and M. Ahmed, "Nursing care systematization with case-based reasoning and artificial intelligence," *Journal of Healthcare Engineering*, vol. 2022, Article ID 1959371, 9 pages, 2022.

[7] S. Abdelmoneem, H. Said, and A. Saad, "Arrhythmia disease classification and mobile based system design," *Journal of Physics: Conference Series*, vol. 1447, Article ID 012014, 2020.

[8] T. Verheyen, A. Decloedt, D. Clercq, P. deprez, S. U. Sys, and G. van Loon, "Electrocardiography in horses-Part 2: how to read the equine ECG," *Vlaams Diergeneeskundig Tijdschrift*, vol. 79, pp. 337–344, 2010.

[9] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajjej, and A. A. Shihab, "Applying dynamic systems to social media by using controlling stability," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4569879, 7 pages, 2022.

[10] E. Grossi and M. Buscema, "Introduction to artificial neural networks," *European Journal of Gastroenterology and Hepatology*, vol. 19, pp. 1046–1054, 2008.

[11] A. Amjad and Al-B. Abdel Rahim, "The possibility of implementing industrial incubators and their role in the development of small industry and medium in Iraq," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–22, 2022.

[12] D.-H. Lee, Y.-T. Kim, and S.-R. Lee, "Shallow landslide susceptibility models based on artificial neural networks considering the factor selection method and various nonlinear activation functions," *Remote Sensing*, vol. 12, p. 1194, 2020.

[13] S. Xu, M.-W. Mak, and C.-C. Cheung, "Deep Neural Networks versus Support Vector Machines for ECG Arrhythmia Classification," in *Proceedings of the 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 127–132, Hong Kong, July 2017.

[14] M. B. Alazzam, A. T. Al-Radaideh, R. A. Alhamarnah, F. Alassery, F. Hajjej, and A. Halasa, "A survey research on the willingness of gynecologists to employ mobile health applications," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1220374, 2021.

[15] J. Sharma, V. Kumar, S. Ayub, and J. Saini, "Uniform sampling of ECG waveform of MIT-BIH normal sinus rhythm database at desired intervals," *International Journal of Computer Application*, vol. 50, pp. 6–9, 2012.

[16] M. Fadhil Jwaid, "An efficient technique for image forgery detection using local binary pattern (hessian and center symmetric) and transformation method," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–11, 2022.

[17] M. Bader Alazzam, H. Mansour, M. M. Hammam et al., "Machine learning of medical applications involving complicated proteins and genetic measurements," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1094054, 2021.

[18] S. Mondal and P. Choudhary, "Detection of normal and abnormal ECG signal using ANN: selected revised papers from the joint," in *Proceedings of the International Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2017)*, pp. 25–37, Pattaya, Thailand, December 2018.

[19] P. Dahlberg, U.-B. Diamant, T. Gilljam, and A. Rydberg, "QT correction using Bazett's formula remains preferable in long QT syndrome type 1 and 2," *Annals of Noninvasive Electrocardiology*, vol. 26, 2020.

[20] M. Pooyan and F. Akhoondi, "Providing an efficient algorithm for finding R peaks in ECG signals and detecting ventricular abnormalities with morphological features," *Journal of Medical Signals and Sensors*, vol. 6, pp. 218–223, 2016.

[21] Al-A. Ishaq Yousef and R. S. Abdul, "Social intelligence and its relationship to decision quality," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–22, 2022.

[22] A. S. Al-Obeidi, N. A. Sultan, and A. R. Obaid, "The degree of applying electronic learning in the Gifted School/Nineveh in Iraq and what management provided to the students and its relationship to qualitative education under coronavirus (COVID-19) pandemic," *International Journal of Computer Applications in Technology*, vol. 66, no. 3-4, pp. 286–293, 2021.

[23] M. Rasmi, M. B. Alazzam, M. K. Alsmadi, I. A. Almarashdeh, R. A. Alkhasawneh, and S. Alsmadi, "Healthcare professionals' acceptance Electronic Health Records system: critical literature review (Jordan case study)," *International Journal of Healthcare Management*, vol. 13, no. sup1, pp. 48–60, 2020, þ.

[24] P. Werbos, "Intelligence in the brain: a theory of how it works and how to build it," *Neural Networks*, vol. 22, pp. 200–212, 2009.

[25] H. An, Md A. Ehsan, Z. Zhou, and Y. Yi, "Electrical Modeling and Analysis of 3D Synaptic Array Using Vertical RRAM Structure," in *Proceedings of the 2017 18th International Symposium on Quality Electronic Design (ISQED)*, pp. 1–6, Santa Clara, CA, USA, March 2017.

[26] J. Zupan, "Introduction to artificial neural network (ANN) methods: what they are and how to use them," *Acta Chimica Slovenica*, vol. 41, 1994.

[27] G. Kumar et al., "Arrhythmia Detection in ECG Signals Using a Multilayer Perceptron Network," AICS, 2019, http://ceur-ws.org/Vol-2563/aics_33.pdf.

[28] W. EnbeyleA. S. Al-Obeidi et al., "Trend analysis and prediction on water consumption in southwestern Ethiopia," *Journal of Nanomaterials*, vol. 2022, Article ID 3294954, 2022.

Research Article

Design of Table Tennis Training Competition Knowledge Interaction Platform Integrating Improved Swarm Intelligence Algorithm

Deqi Li [ID]

Department of Physical Education, Zhejiang Yuexiu University of Foreign Languages, Shaoxing, Zhejiang 312000, China

Correspondence should be addressed to Deqi Li; 20011018@zyufl.edu.cn

Table tennis is China 's national game and the proudest sport in China's sports field. During the research and technology service work of the Chinese table tennis team for many years, it has accumulated a large amount of valuable data on the analysis of skills and tactics of training and matches, match video, training monitoring, and so on. This paper discusses the relevant theory of swarm intelligence algorithm processing big data on the table tennis training competition knowledge interaction platform system, as well as the technical support of Nginx and Tomcat, and determines the technical basis of the table tennis training competition knowledge interaction platform. Through the establishment of the firefly algorithm model, the resource search ability is enhanced, and the traditional firefly algorithm is improved. From the results of the system performance test, it can be found that the improved swarm intelligence algorithm adopted in this paper improves the global convergence, and the load balancing degree gradually decreases with the increase of time. The improved firefly algorithm shows good performance when the bandwidth is low, and the resource occupancy rate is greatly reduced. When the bandwidth is 20, it is reduced by 12.55%. It solves the shortcomings of long time and low success rate, so as to verify the convenience of the system operation and the power of functions and make the platform more intelligent and efficient.

1. Introduction

For more than half a century, the Chinese table tennis team has gone through hardships and struggled tenaciously, winning the world table tennis championship time and time again, creating the miracle of Chinese table tennis, occupying a core position in the hearts of the Chinese. It is the pride of the Chinese. This is an aperiodic exercise based on speed, power and flexibility, and aerobic metabolism. It is an organic combination of skill, physical fitness, and intelligence and is characterized by fast speed, many turns, and great strength. Athletes must possess a wide range of physical qualities, including speed, dexterity and coordination, quick reflexes, and explosiveness. Since the 1950s, researchers have proposed many novel intelligent algorithms. For example, the artificial neural network algorithm that abstracts the human brain neuron network from the

perspective of information processing, the ant colony algorithm derived from the behavior of ants searching for food to generate a path, and the particle swarm algorithm that iteratively seeks the optimal solution by imitating a swarm of birds or fish are widely used in processing a deeper study of big data analysis and algorithm theory.

Sports teams and research groups collect vast amounts of data during daily training and competition, including game data, training tracking data, sports injury data, physiological and biochemical tests, technical and tactical analysis, and other types of information. The cloud data center is responsible for transmitting these data, and the efficiency and reliability of data transmission will directly affect the performance of the cloud data center. A data migration strategy is the basic requirement for data migration, and it is also a solid guarantee for a stable and efficient system in the future. A good migration strategy can not only reduce migration

costs, but also improve cloud data center maintenance and management. Currently, there is no uniform and efficient way to manage this information. How to make full use of these data to provide a platform for coaches and athletes to acquire knowledge and process information and help athletes improve their training levels and achieve good results in various competitions is a problem that needs to be solved.

The focus of this research is on the construction of an interactive platform for table tennis team practice and competition, integrating data in various formats, such as video, image, file, and chart formats. It can provide a simple and efficient interactive interface for table tennis coaches, athletes, scientific researchers, medical staff, managers, etc., to facilitate their acquisition of relevant information, so that all aspects of training management can be refined and improved, and better services are provided for athletes' training and competition. The platform can manage all kinds of team information in a unified way and provide more scientific and reasonable monitoring, evaluation, and decision-making assistance for coaches and athletes' training and competition, so as to help athletes achieve better results.

2. Related Work

As things stand, table tennis can be said to be unsurpassed by any sport in China. An essential factor affecting the system design of table tennis robot for automatic detection of table tennis skill and metrics is the spin condition, path, and bouncing power of table tennis. However, the common prediction algorithms could not cope with the data of time sequence and the corresponding spinning state. Zhang researched the methodology of compensated vague neural network-based trajectory prediction for table tennis technical and tactical index automatic detection. Test results showed that the convergence time of the compensated fuzzy neural network was shortened, the training time was shortened, and the prediction accuracy was improved [1].

Deep mining plays an increasingly important role in the formulation of tactics and rules of sports events. Based on the data mining platform, Choi analyzed the correlation between the dynamic score of table tennis and the game performance. Data mining techniques can extract a superset of information from a variety of data sources and then combine this information, so that some patterns and internal relationships can be found. The results showed that the proportion of score reversals gradually decreased as the table tennis game progressed. This was mainly due to physical and psychological factors. Through the analysis of big data, the imbalance between the total score and the game result was explored. Therefore, athletes should carry out targeted combination exercises to reduce the lack of adaptation in the game [2].

Along with the application of computer large data, sports large scale data profiling offers scientific reference for the improvement of athletes' training mode. Based on big data analysis, Zheng analyzed the significance and usage of mental element coaching for table tennis education. In the actual game, psychological quality has become an important foundation for athletes' ability and skills. The usage of the

mental factor training in table tennis should make students pay attention to their own characteristics and exercise mental quality on this basis. At the same time, competitive motivation is the main factor affecting the formation of personal emotional changes. Teachers should help players establish correct table tennis motivation [3].

Yilmaz et al.'s study aimed to investigate the effects of the interactive memory system (TMS) and interactive deck in collaborative computer-supported learning (CSCL) on learners' perception of having a social existence and ability to self-regulate. The perceived social presence and self-regulation skills of students in collaborative groups were compared while constructing knowledge on wiki, blog, podcast, and Facebook platforms during CSCL within the scope of the study. When examining the effect of interactive platforms and TMS on self-regulation in terms of self-regulation skills, it was observed that, despite the significant effect of interactive platforms, the effect of the combination of TMS individually and interactive platforms TMS was not significant [4].

With the development of time, data and video information about table tennis training and competitions are increasing day by day. As is known to all, scientific material selection, scientific training, and scientific management are closely related to the improvement of competitive sports level. The fast evolution of technologies in science needs to push the pace of scientific management of sports, for which some research work on swarm intelligence algorithms has been collected.

Recently, group intelligence-based algorithms have become an emerging area of research. Of these, effective solutions to different complex problems have been found with two superheuristic algorithms that were developed recently, that is, the Cuckoo Search Algorithm (CSA) and the Firefly Algorithm (FA). Bhattacharjee and Sarmah attempted to use them to solve combinatorial problems, especially the 01 Knapsack Problem (KP) and the Multidimensional Knapsack Problem (MKP), while in the improved FA, a variable distance shift of the repair operator based on local search and adversarial learning mechanism was used to verify the effectiveness of the algorithm [5].

Ma et al. discussed the algorithm equity of Particle Swarm Optimization (PSO) and a variety of additional newer SI algorithms. Under certain conditions, the primitive releases of SFLA, GSO, FA, ABC, and GSA were identical to the PSO algorithm. Substantiation indicated that the advanced version of ABC performed best on continuous benchmark functions, while the enhanced versions of SFLA and GSA performed best on combined problems with nesting [6].

Ntouni et al. optimized the inspection and testing procedures of spread-based molecular communicative processes by using weighting and detectors that had suitable power transfer values. For this purpose, they proposed a robust iterative optimization algorithm based on standard particle swarm optimization (PSO) techniques, called accelerated-assisted PSO (a-APSO). They implemented the a-APSO operator to estimate the determiner weighing to minimize the error probability in a closed form

representation. The results of the study showed that the margin of error outperformed when using A-APSO weights rather than when using weight values available in the literature or when using weights evaluated by standard PSO algorithms [7].

Mashwani et al. proposed a multi-intelligence algorithm (MSIA) for dealing with bounded constrained functions. The performance of the algorithm was evaluated by using a benchmark function with 30 bounded constraints. The algorithm had good convergence and diversity preservation, approximating promising solutions [8].

To verify the data enhancement results of the characteristic choice alcohol system, which is based on the group smart algorithm, Chen et al. conducted experiments in six classes with similar conditions in the first middle school of the city. The experiments showed that, under the feature selection algorithm, students' academic performance was approximately 30% greater than the other methods of teaching, and the sense of cooperation among students reached 0.8. The contradiction between popularity and improvement was solved, and the problem of differentiation and transformation of late-comers was solved [9].

Gao et al. studied the population smart algorithm for the extraction of spatial facial features and hybrid image element breakdown of hyper-spectral distant images. Firstly, two different spatial spectral feature extraction methods were introduced, then linear and nonlinear spectral hybrid models were studied, and then the end element extraction approach based on hybrid quantum cluster of particle optimization algorithm was investigated. The result of experiments showed that the algorithm improved the population based intelligent algorithm for cluster pixel decomposition. The accuracy of the suggested 3DLBP space spectral features for classification was 94.22%. These swarm intelligence algorithms play a great role in solving multivariate data problems and global optimization, but their practicability is not strong [10]. The above research swarm intelligence algorithm and related content of table tennis are analyzed in detail. It is undeniable that these studies have greatly promoted the development of the corresponding fields. Much can be learned from methodology and data analysis. However, there are relatively few studies on table tennis motion by swarm intelligence algorithms, and it is necessary to fully apply these algorithms to the research in this field. The approaches offer a few points of reference for this study, but because of the brief duration and low sample volume of the associated research, this investigation has not yet gained public acceptance.

3. Algorithm and Integration

3.1. Knowledge Interaction Platform for Table Tennis Training and Competition. With the vigorous development of modern sports, the data about table tennis training and competitions is increasing rapidly, and it is impossible to comprehensively obtain beneficial experience for coaches and athletes only by manually collecting and analyzing the data of previous competitions. In view of the characteristics of table tennis knowledge interaction platform with complex
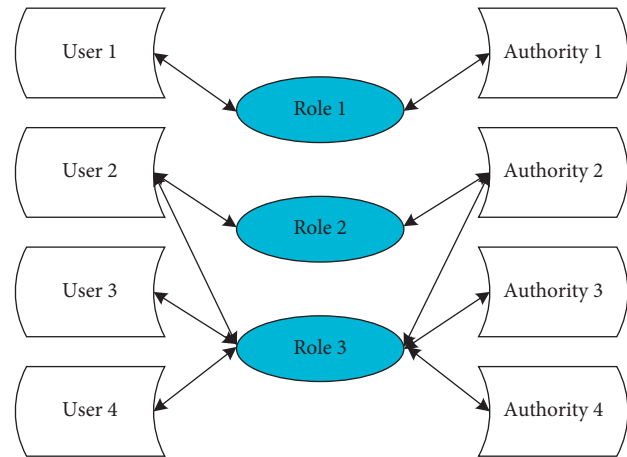

Figure 1: RBAC fundamentals.

data volume, many user types, and large number of users, role-based access control (RBAC) has been introduced [11]. Its basic idea is to assign access rights to roles and then assign roles to users, and users can only get appropriate access rights through roles. A user can have multiple roles, and the same role can be assigned to multiple users, so there is a many-to-many relationship between users and roles. Users can dynamically activate the roles they own according to their own needs, so they have all the rights to activate the corresponding roles [12]. In RBAC, a user's authority is granted and revoked by assigning and revoking roles, which makes the user's authority and access logically separated, which greatly simplifies authority management [13]. Figure 1 shows the relationship among users, roles, and permissions in the basic principles of RBAC.

Table tennis knowledge interaction platform is a comprehensive management platform. On the one hand, it needs to manage a large amount of data, and on the other hand, it needs to provide a friendly interface for specific users. Therefore, it is divided into two systems: one is a front-end system presented to coaches, athletes, and researchers in the form of a website, providing various data query, professional analysis, retrieval, management, and other functions [14]. And the other is a data input system that provides various data, providing functions such as input, deletion, and management of various data. The data entry system provides functions such as inserting, deleting, and managing various data. The architecture adopts two modes of browser/server (B/S) and client/server (C/S), leverages the strengths of both architectures to deliver coaching, athletes, and researchers with a simple and friendly interface, and ensures the security of the back-end data input system [15]. This is illustrated in Figure 2.

According to the current situation of national table tennis team training, competition and team management, the platform is divided into 6 subcolumns: athlete database, event database, monitoring database, business learning database, team database, and user management, as shown in Figure 3. The tracking database is the core part of the platform and is divided into 5 subsections of training, fitness, form and function, education, and competition monitoring.
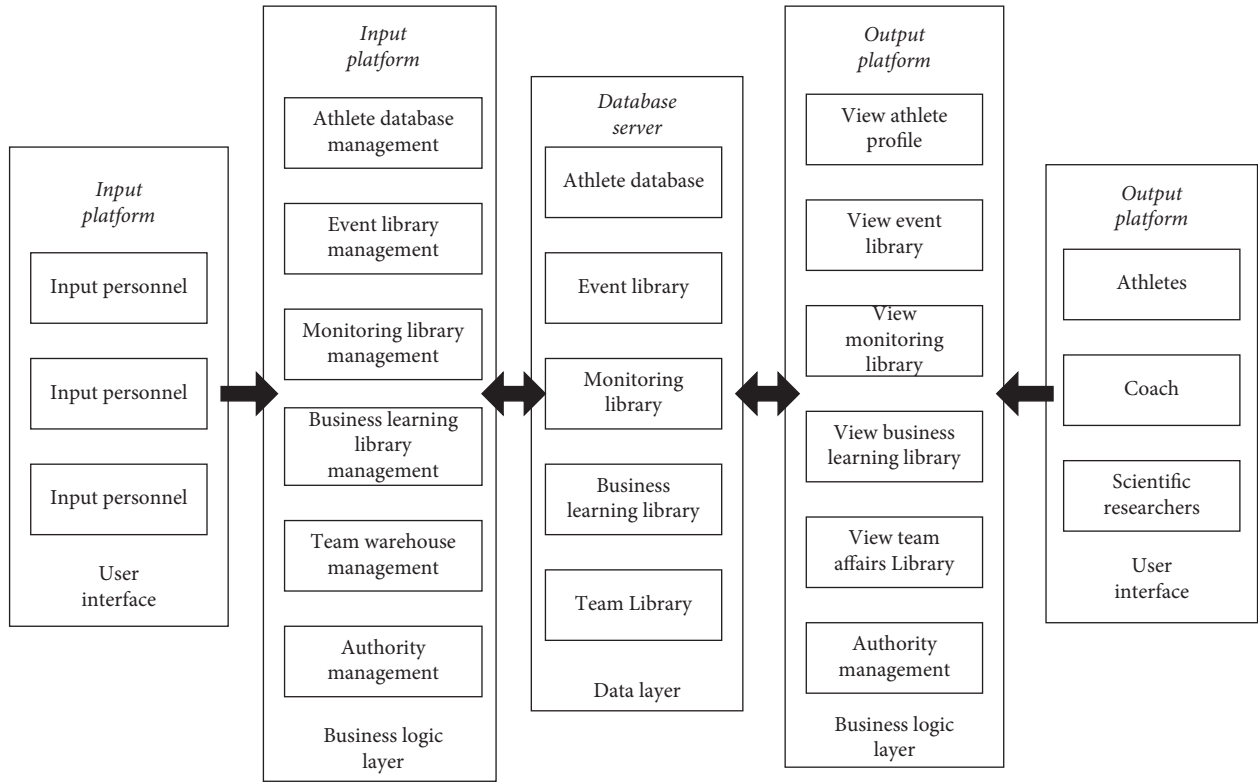
Figure 2: Platform structure design diagram.
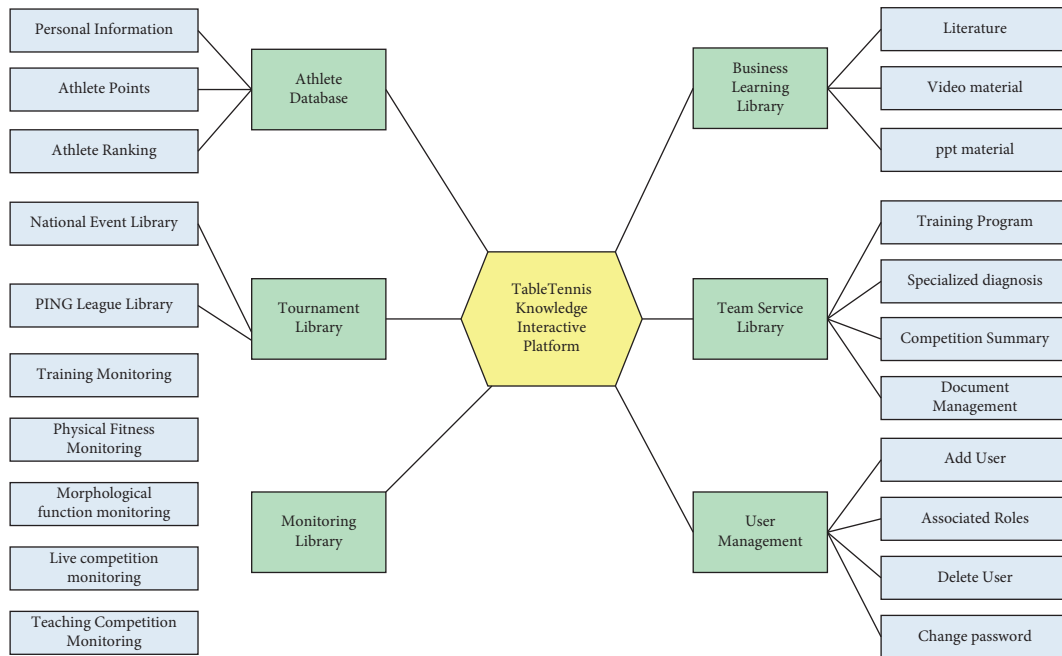
Figure 3: Functional diagram of table tennis team training and competition knowledge interaction platform.

By monitoring the technical and tactical training level of athletes in routine training or competition training, it is convenient to understand the development trend of athletes' training level.

The Business Learning Library provides literature, videos, and PPT material for technical and tactical analysis.

Table tennis literature is divided into many categories, with distinct themes and various forms, including many game videos provided by major domestic and foreign video organizations such as the International Table Tennis Federation and China Network Television, as well as PPT data on the national table tennis team's skill and strategy analysis of

domestic and international players during the tournament. A knowledge search engine and video library about table tennis was established for the first time. The scientific information contained in this module is important for improving the cognitive culture and coaching skills of coaches, as well as the coaching skills and athletic performance of players.

3.2. Swarm Intelligence Algorithm.

Swarm intelligence originates from the research on the group behavior of social insects represented by ants and bees. Swarm intelligence algorithm is a computing technology that evolves according to group behavior. There are many examples of swarm intelligence in nature, such as bird migration, ant colony foraging, fish gathering, and bacterial growth. These gregarious organisms with low individual intelligence behavior gather in ecosystems and adapt to transforming natural environment. Taking ants as an example, Figure 4 shows the foraging behavior of ants. It can be seen that the ants show amazing intelligent behavior in the process of foraging. A is the ant cave, D is the food source, and EF is the obstacle. As shown in Figure 4(b), after a period of time, the information volume of the path BEC is 1/2 of the BFC. After another period of time, 20 ants go from points B, F, and C to food source D, as shown in Figure 4(c). In the following time, more and more ants will choose BFC, and eventually all ants will select BFC. The closest routes among ant burrows and food sources are found by means of ant colonial optimization.

The behavioral norms that should be followed in the swarm intelligence system include several principles: the principle of adaptability, the principle of neighbors, the principle of response diversity, the principle of quality, and the principle of stability. The principle of stability expresses that a group should not change its behavior every time the environment changes. That is to say, individuals should have the ability to adapt to the environment, and when the external environment changes, they should adjust their behavior as soon as possible; individuals should carry out simple information exchange and transmission; group activities should show diversity. There are quality factors that influence each other.

There is no central control in the swarm intelligence algorithm, so when one or several individuals perform poorly, it will not affect the solution of the problem to the entire group. The advantages of swarm intelligence are one of the major reasons why it has become a research hotspot. It has gradually developed into as one of the top priority research topics in the field of Smart Computing. As the research progresses, more and more swarm intelligence algorithms are proposed and applied to more fields; especially the ant colonies intelligence theory of swarm optimization algorithm, particle swarm optimization algorithm, firefly optimization algorithm, etc. have been successfully applied to many fields.

This paper proposed a method, that is, firefly algorithm of improved bandwidth (IB-FA), based on the Firefly algorithm to solve the data migration problem in the cloud computing environment. The IB-FA algorithm adopts a heuristic location selection and placement strategy for the multiobjective constrained optimization problem to realize the data migration problem. In the firefly algorithm process of IB-FA, each server is used to generate a population in the form of firefly individuals, and an objective function that can describe the resource utilization, migration cost, and network bandwidth in the cloud computing environment is designed as the fitness function of the firefly algorithm process in IB-FA, which is used to calculate the absolute brightness of the individual in the firefly algorithm. An individual with a large absolute brightness value can attract an individual with a small absolute brightness value to move to it, and the convergence of the optimal solution can be achieved by an iterative method. However, in the later stage of the classical firefly algorithm, many individuals with relatively small fitness function values will gather near fireflies with better fitness function values, and the distance between individuals gradually shrinks, so the local search ability is weakened accordingly. In response to this problem, the IB-FA algorithm introduces adaptive inertia weights to avoid individual fireflies falling into local optimum. IB-FA obtains the global best problem, which is the final solution when the iteration counts reach the maximum. The improved firefly algorithm can save the bandwidth of the cloud computing data center, improve the load situation of the data center, and, at the same time, help improve the user experience and better realize the optimal location selection strategy for data migration.

3.3. Mathematical Description and Analysis of IB-FA.

Construct p feasible solutions $x_1, x_2, x_3, \ldots, x_p$ in the solution space to represent m target data center nodes to be selected, and then, the parameters involved are as follows: the number of fireflies is p, r_{ij} is the distance between fireflies i and j, the absorption coefficient is λ, the absolute brightness is I_i of firefly i, and the attractiveness is α_{ij}.

(1) The initial population $X^g = \left\{ X_1^g, X_2^g, \ldots, X_P^g \right\}$ of fireflies is initialized, where X_P^g represents the position of the ith firefly in the g-th generation, $i \in [1, m]$.

(2) The relationship between the absolute brightness I_i of firefly i and the objective function value is established. Usually, the objective function is used to describe the absolute brightness. The better the objective function is, the larger the absolute brightness value is; that is,

$$I_i = h\left(X_i^g\right). \tag{1}$$

$h\left(X_i^g\right)$ is defined as

$$h\left(X_i^g\right) = \text{acrDC}_i - v_j^i. \tag{2}$$

acrDC_i represents the available resources of the ith server, v_j^i represents the migration cost from server i to server j, μ represents the migration cost per minute, t_j^i represents the time required for migration, and v_j^i is defined as
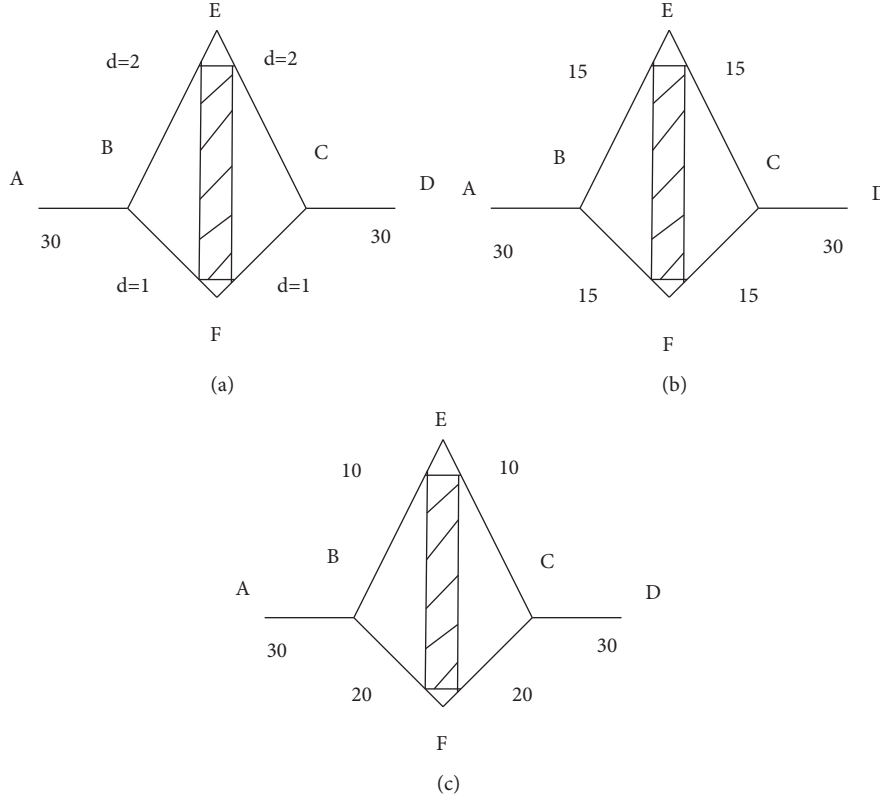
(a)

(b)

(c)

FIGURE 4: Intelligent behavior of ants during foraging. (a) The two paths selected by A–D. (b) Variation of pheromones in two paths. (c) Final path selection.

$$v^i_j = \mu \times t^i_j. \tag{3}$$

(3) Since the traditional attraction calculation only considers the distance between fireflies and the maximum attraction, it does not take into account the absolute brightness of fireflies, which is inconsistent with the actual situation. The method of attraction α_{ij} is defined. α_{ij} is defined with regards to the difference between the two fireflies and the official brilliance of the fireflies. If $I_i > I_j$, then the attraction between two fireflies is defined as

$$\alpha_{ij} = I_i \times \exp\left(-\lambda R^2_{ij}\right). \tag{4}$$

(4) Update the position of fireflies according to the attraction among fireflies. Firefly j is drawn to firefly i and travels toward firefly i. Then, the equation for updating the position of firefly j is defined as

$$X_j(T+1) = \omega X_j(T) + \beta_{ij}\left(X_i(T) - X_j(T)\right)$$
$$+ \beta\left(\text{rand} - \frac{1}{2}\right). \tag{5}$$

Among them, $X_i(T)$ represents the spatial position of firefly i in iterations T. β represents the step factor, which is a constant in [0, 1]. In order to avoid the fireflies oscillating back and forth at or near the local extreme point, an adaptive

inertia weight ω is added, which decreases linearly with the increase of the number of iterations and the objective function, which is beneficial to local search. So, the definition of ω is as follows:

$$\omega(T) = f\left(X^g_i\right) \times \exp(-T). \tag{6}$$

3.4. Formalization of the IB-FA Algorithm.

The problem can be formalized as the problem of migrating q pieces of data to p servers in a cloud data center. The solution can be represented as an n-dimensional vector, where each element value represents the location of the destination server node of the data to be migrated. Assuming that each server node will dynamically change its state according to the load, the impact of different location migration strategies on the data center network bandwidth is also different. A quad is defined as follows:

$$X = \{D_i, \text{DC}, B, \text{BS}\}. \tag{7}$$

D_i represents the set of q data to be migrated in the ith server, denoted as

$$D_i(n, \gamma) = \left(D_{i1}, D_{i2}, \ldots, D_{iq}\right), \quad i \in [1, m]. \tag{8}$$

γ indicates the time point, and D_{ij} indicates the j th data to be migrated in the ith server. DC represents the set of p available server nodes, denoted as

$$\mathrm{DC}(m, \gamma) = \{\mathrm{DC}_1, \mathrm{DC}_2, \ldots, \mathrm{DC}_p\}. \tag{9}$$

$B(m, \gamma)$ represents the network available bandwidth of m servers at γ time, denoted as

$$B(m, \gamma) = \{B_1, B_2, \ldots, B_p\}. \tag{10}$$

The primary goal of solving the problem in this paper is to find a set of candidate server locations that maximize performance and minimize migration cost, denoted as

$$\mathrm{BS}\big[D_{ij}, \mathrm{DC}_m, v_m^j, \gamma\big], \quad (i \neq m). \tag{11}$$

Then, the location node that saves the most bandwidth in the entire data center is selected. S represents the sum of the sizes of q data to be migrated, denoted as

$$S = \sum_{i=1}^{q} S(D_{ij}). \tag{12}$$

T represents the time required for data transmission, and the calculation formula is defined as

$$T = \frac{S}{\mathrm{Bandwidth}}. \tag{13}$$

η indicates the bandwidth usage of the current server, defined as

$$\eta = \frac{T_w}{T_w + \sin\theta T_b} \times 100\%. \tag{14}$$

Since the execution of the background data of each server is constantly changing, the corresponding execution time is also constantly changing. $\sin\theta$ in formula (14) is used to simulate the variation factor $\theta \in (0, \pi)$ of this background data. It can be seen from formula (14) that if the value of η is smaller, the server bandwidth occupancy rate is smaller; then, for the bandwidth of the entire data center, there is more bandwidth that can be used to process other transactions, which is conducive to improving the bandwidth utilization of the entire data center.

Niche is a widespread phenomenon in the evolution theory of nature. Biological individuals usually tend to live in the same environment as individuals similar to themselves. It is precisely because of this that nature is full of vitality. Niche originates from the biological idea of a niche, where organisms of the same species live together to create small habitats that, in turn, allow different types of individuals to inhabit them. Applying this idea to computational science means that data with the same trait will be grouped, and data with different traits will be separated, thereby avoiding large datasets around local optimum points and increasing population diversity. The main goal of a niche algorithm is to establish and maintain a stable, diverse subpopulation that can perform parallel evolutionary searches in different research fields, thus solving the problem of optimizing multiple peaks and targets.

Like other smart algorithms, the standard firefly algorithm tends to converge on perfect sites and mature prematurely, improves the efficiency of stereo mapping by using stereo mapping to obtain initialized firefly populations, and

provides a basis for population diversity. It uses random momentum weights to change the position update pattern of individual fireflies, striking a better balance between investigative and useable. The timely implementation of niche technologies can not only eliminate the phenomenon of falling into local extremes, but also increase the diversity of populations. The firefly algorithm suffers from falling into local extremes and reducing population diversity at the end of the iteration, so the advantages of the niche technique are used to offset the shortcomings of the basic firefly algorithm.

In general, the initial distribution of individual firefly positions is random, which can lead to an uneven distribution of firefly positions. Therefore, this chapter introduces a chaotic representation with normal and ergodic features to obtain a chaotic initialized population, which not only prevents the algorithm from falling into local optima, but also increases the diversity of the population. Due to the better sequence homogeneity, the cube representation was chosen. The phases of chaos initialization are as follows:

(1) The first individual is obtained. The first firefly is represented by a randomly generated D-dimensional vector, $R = r_1, r_2, \ldots, r_d$, where $r_i \in [-1, 1]$, $1 \leq i \leq d$.

(2) The remaining $N-1$ individuals are obtained. Perform $N-1$ iterations for each dimension equation of R:

$$r(n-1) = 4r(n)^3 - 3r(n). \tag{15}$$

Among them, $-1 \leq r(n) \leq 1$, $n = 0, 1, 2, \ldots$.

(3) The chaotic variables are mapped to the solution space with the following formula:

$$x_{id} = O_d + (1 + r_{id}) \frac{(R_d - O_d)}{2}. \tag{16}$$

Among them, x_{id} represents the dth dimension coordinate of the ith firefly in the solution space; O_d represents the lower limit of the dth dimension of the solution space; r_{id} represents the dth dimension of the ith individual obtained by formula (16); R_d represents the upper limit of the dth dimension of the solution space.

If an individual is closer to the optimal individual, a smaller weight will be generated, thereby speeding up the convergence; if a larger value is generated, a larger fitness function value will be obtained, which is worse than the optimal value; these weights are eliminated. Likewise, the situation is basically the same when the individual is far from the optimal solution. Therefore, in this paper, random inertia weights are introduced into the position update formula, which is

$$\begin{aligned} x_i(t+1) &= v * x_i(t) + \beta(x_i(t) - x_i(t)) + \alpha\theta_i, \\ v &= \mu_{\min} + (\mu_{\max} - \mu_{\min}) \times \mathrm{rand} + \lambda \times \mathrm{rand}n. \end{aligned} \tag{17}$$

$\mu_{\min}$ and $\mu_{\max}$ are the minimum and maximum weights, respectively; rand are the uniformly spread random numbers in the region $[0, 1]$. λ is the measure of the deviation level between v and its mean, which is used to correct the weight

error of the obtained values to make the weight change towards the ideal weight because, in general, the error of the experimental results follows a normal distribution. rand$n$ is a normally distributed random number.

At the beginning of the iteration, the basic firefly algorithm performs a global search, and there is no need to introduce the niche technology. When firefly populations are highly aggregated, this is an opportune time to introduce niche technologies. When individuals are grouped, it can be seen that the fitness function values of most individuals are relatively close, so the degree of grouping between individuals should be expressed by defining variables, as shown in

$$\phi^2 = -\sum_{i=1}^{p}\left(\frac{f_i - f_{\text{avg}}}{f}\right)^2. \tag{18}$$

Here, p is the number of individuals, f_i is the value of the fitness function for the ith individual, and f_{avg} is the current average fitness of the firefly population. There is a sufficiently small threshold T; when $\phi^2 < T$, firefly populations are highly aggregated, which is an appropriate time to introduce niche technology. This in turn leads to the exclusion of unsuitable individuals, thereby maintaining population diversity.

3.5. Design of Knowledge Interaction Platform for Table Tennis Training and Competition. The knowledge interaction platform system for table tennis training and competition uses the Java language and adopts the J2EE architecture of JDK1.8. The page development uses the powerful EasyUi framework, and the background is built with the SpringMVC framework. The development tool used in this system is Myeclipse2014 or Myeclipse2010. The software development environment is the software required for the development system and its version requirements. The configuration details are shown in Table 1.

Online interactive system has been relatively stable in technology and is in a high development stage. At the same time of development, it also develops a variety of platform technologies. The online platform architecture that is widely accepted and adopted by various user units is the B/S three-tier architecture, as shown in Figure 5.

From Figure 5, a general online system has three layers: front-end server tier, implementation server tier, and database tier. The front-end server tier is to deliver a web-based interface for users. As a learning portal, the main use method is that users log into the front-end web server with a browser, and the server will provide services to users in the form of web pages. The application server layer is the layer that provides the data server layer operation interface for the front-end server layer. The data server layer is the core of providing data services for the entire system. When a user accesses the online learning system through a personal computer or mobile device and performs various operations, the front-end server layer first parses and decomposes the user's operation and then distributes it to the application server. According to the user request forwarded by the front-end server, the application server layer calls the data server layer for service, obtains the read or write information requested by the user, and feeds back the result to the user. The advantages of establishing an online learning system for users in this way are that a common architecture platform can be used, the technology is relatively mature, and an appropriate content platform can be designed according to their own conditions.

Tomcat is one of the most popular web application servers in the world and is well respected by developers and Java developers. Tomcat servers can be unstable and sometimes fail to run due to heavy network traffic, while Nginx can handle large numbers of requests and large amounts of data consistently and efficiently with low resources.

Nginx is a high-performance HTTP server and reverse proxy, which can manage static resources well on multiple high-load websites, separate dynamic and static pages according to programming rules, and use URL hashing, iphash, polling, etc. to balance the load of internal servers. Figure 6 shows the dynamic and static architecture of the combination of Tomcat and Nginx. Nginx is placed in front of Tomcat as a reverse proxy, so that Nginx can provide static resources such as pictures, js, css, and html.

4. Table Tennis Training Competition Knowledge Interactive Platform Simulation Experiment

The main purpose of system performance testing is to test the load capacity and responsiveness of the network and to check the compatibility of the platform's function with the design requirements. The content of performance testing includes the following:

(1) Response speed when many users concurrently have access.

(2) Resources of the system in case of concurrent access by a large group of users.

(3) Network transmission efficiency and stability during concurrent visits by many users.

Figure 7 shows the convergence curves of the model Firefly algorithm and the improved algorithm. From Figure 7, the improved algorithm has better convergence performance. Table 2 shows the comparison between the firefly algorithm and the improved firefly algorithm. Based on the performance comparison data in Table 2, it can be concluded that the improved hybrid algorithm has better performance in all aspects and shortens the running time of the algorithm.

To validate the efficiency of the improved Firefly algorithm for target location selection policy load balancing of migrated data during cloud data center operation and to compare the bandwidth usage, data values were collected every two weeks for a testing period of 10 weeks. As shown in Figure 8, as the cloud data center operation time increases, the load balancing degree of the two migration strategies gradually decreases. The improved firefly algorithm reduces the load balancing degree of the firefly algorithm. The firefly

TABLE 1: Software development environment.

Environment	Configuration requirements	Environment	Configuration requirements
Operating system	Windows 10	Browser	IE9 and above or Firefox, Google
Database server	Sql server 2008 or above	Java development environment	Jude v6.1
Application server	Tomcat 8.0	UML modeling tools	Jude v6.1
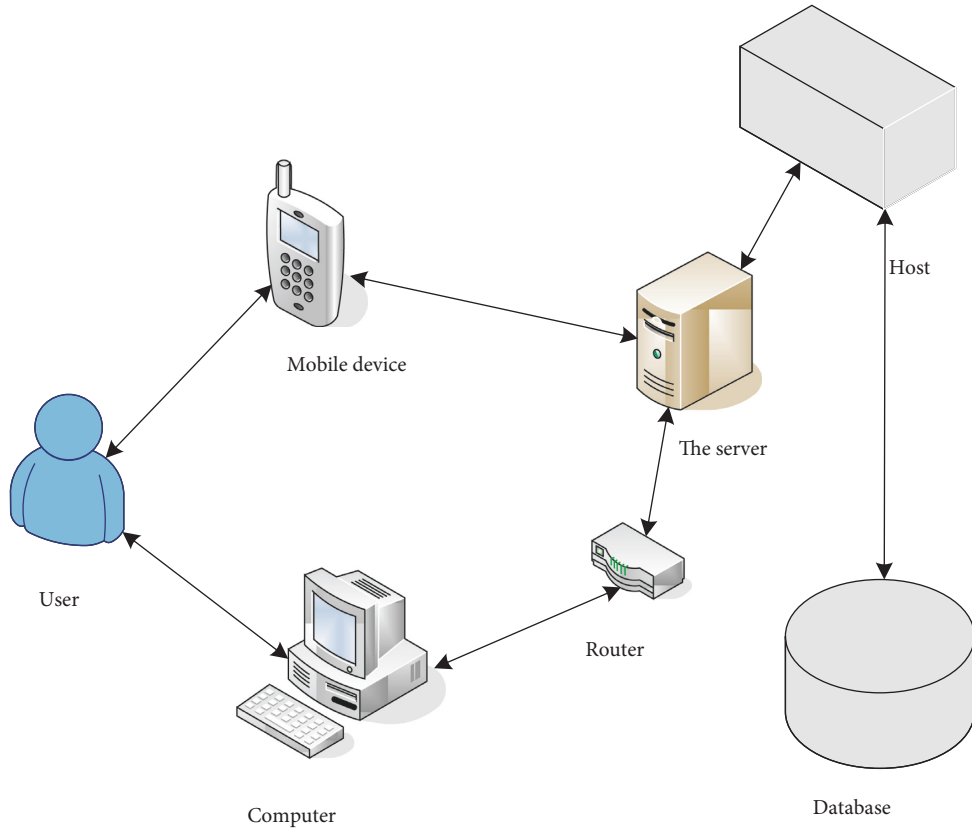
FIGURE 5: B/S three-layer structure of a project.
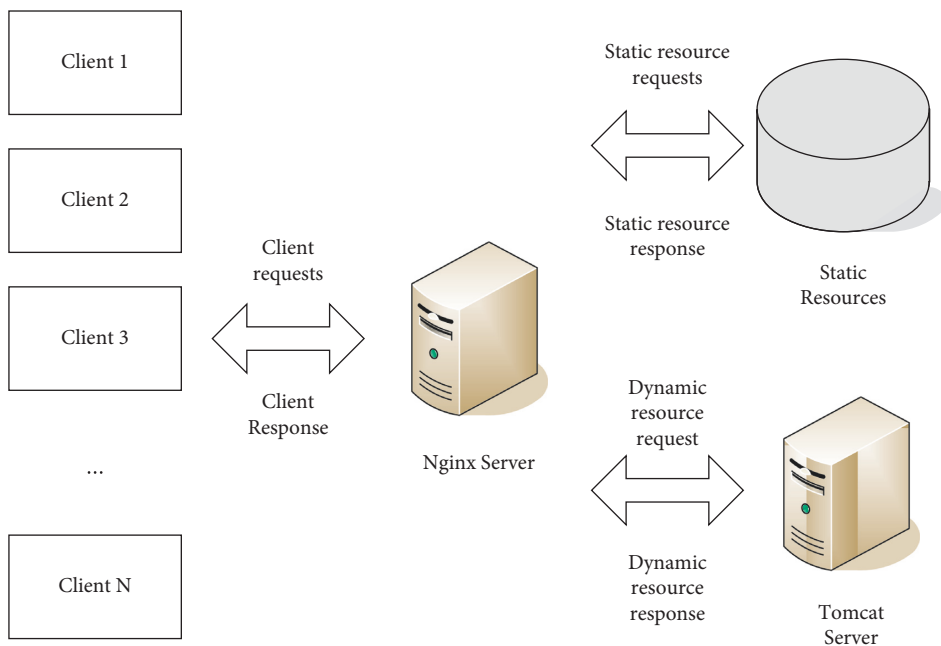
FIGURE 6: Dynamic and static separation architecture of Tomcat combined with Nginx.
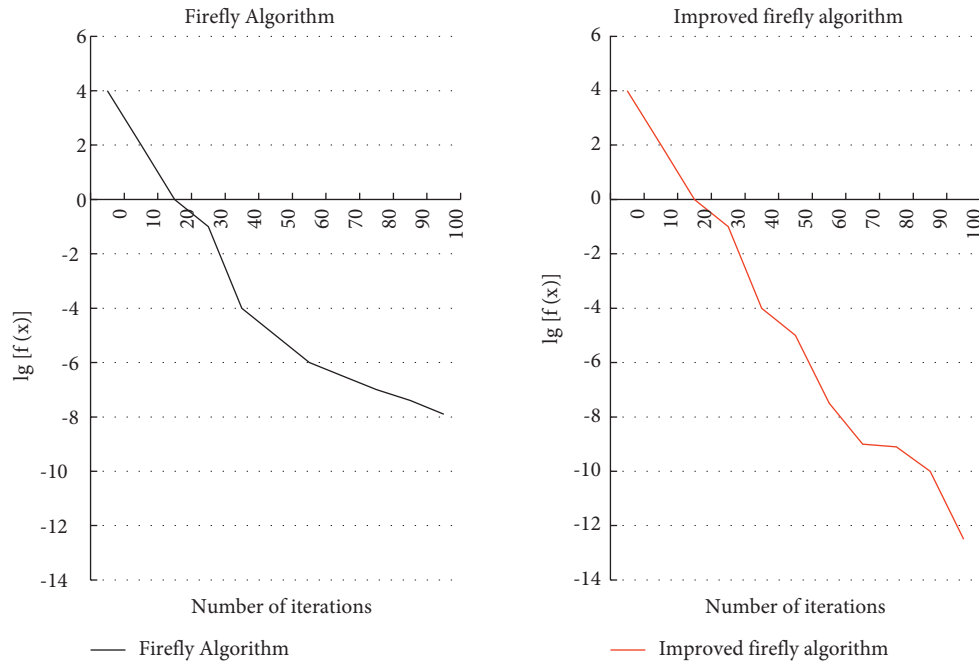
FIGURE 7: Concentration curves of the firefly algorithm and the modified firefly algorithm.

TABLE 2: Comparison of the firefly algorithm and the modified firefly algorithm.

Performance indicators	Firefly algorithm	Improved firefly algorithm
Running time	46.31 s	33.52 s
Average value	$2.4121e-7$	$5.3247e-9$
Optimal value	$3.5214e-7$	$1.2457e-9$
Worst value	$6.6485e-6$	$2.6458e-8$
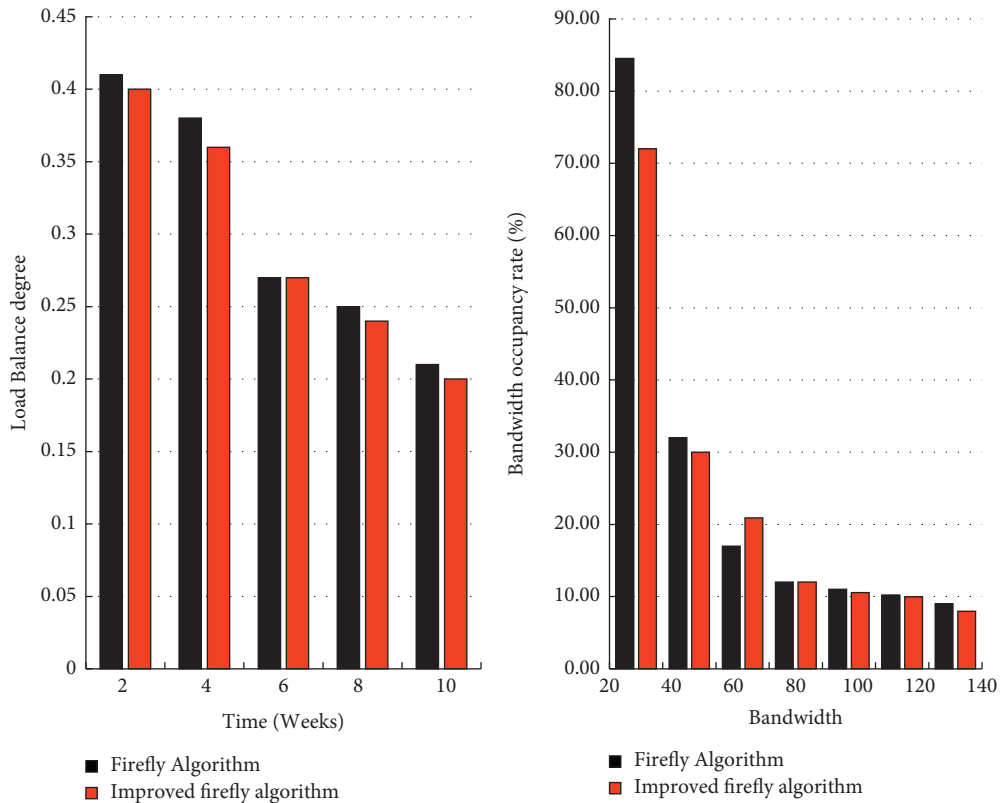
FIGURE 8: Load balancing comparison chart and comparison chart in terms of bandwidth usage.
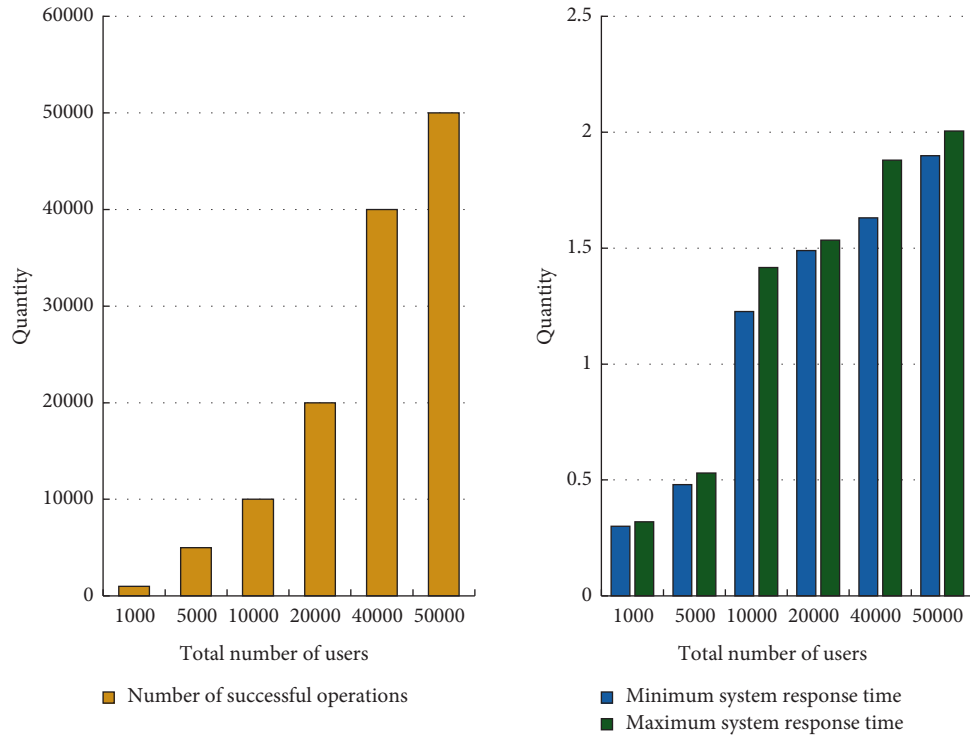
FIGURE 9: System performance test results.

policy finds the migration location expeditiously. Though it relieves the pressure of burden on the data location pending migratory data, it does not consider the load capacity of the target location in detail, which results in sometimes good load effect and sometimes bad effect. As the bandwidth gradually increases, the ability to process each service becomes stronger and stronger, and the corresponding bandwidth occupancy rate gradually decreases. The firefly policy has considered the resource occupation of the destination position, and the chosen destination position can achieve improved mobility and ensure smoothness. However, compared with the improved firefly algorithm, the effect is not very ideal. The improved firefly algorithm has a higher bandwidth. When the bandwidth is low, it shows good performance, and the resource occupancy rate is greatly reduced. When the bandwidth is 20, it is reduced by 12.55%. Experiments showed that the improved Firefly migration strategy performed better in terms of bandwidth usage and allowed more available bandwidth to serve additional services to other servers in the data center in the cloud due to less bandwidth usage.

The system performance test findings for this platform are presented in Figure 9. Table 3 shows the total latency for different number of tasks.

From Figure 9, when the system load reaches 50,000 users, it can still respond normally, fully meeting the performance requirements of the system. The improved firefly algorithm can reduce the total delay of the system to a certain extent under different tasks.

System Function Test: training and game monitoring library is extremely important. From 2018 to 2020, 120

TABLE 3: Total latency under different number of tasks.

Number of tasks	Firefly algorithm of improved bandwidth	Improved firefly algorithm
20	985	976
40	1252	1112
60	1520	1441
80	1630	1530
100	2840	2006
120	3001	2541
140	3060	2733

matches between the top 35 men in the world rankings (based on the ranking in April 2020) and 130 matches between the 21 women's athletes were selected. Among them, there are 70 competitions for Chinese men's athletes and 50 competitions for men's athletes from other countries; 55 competitions for Chinese women's athletes and 75 competitions for women's athletes from other countries. The technical effectiveness (TE value) of the one-three-beat, two-four-beat, and stalemate ball in each game is calculated, and the percentile method is used to build a technical performance evaluation model for excellent table tennis players. TE $\geq 70\%$ is excellent, and TE $< 30\%$ is bad.

After inspection, the data of the athletes' technical performance TE values are all in a normal distribution. According to the technical performance TE value evaluation method, the evaluation grade standards for the technical performance TE values of outstanding male and female table tennis players have been obtained, as shown in Figure 10.
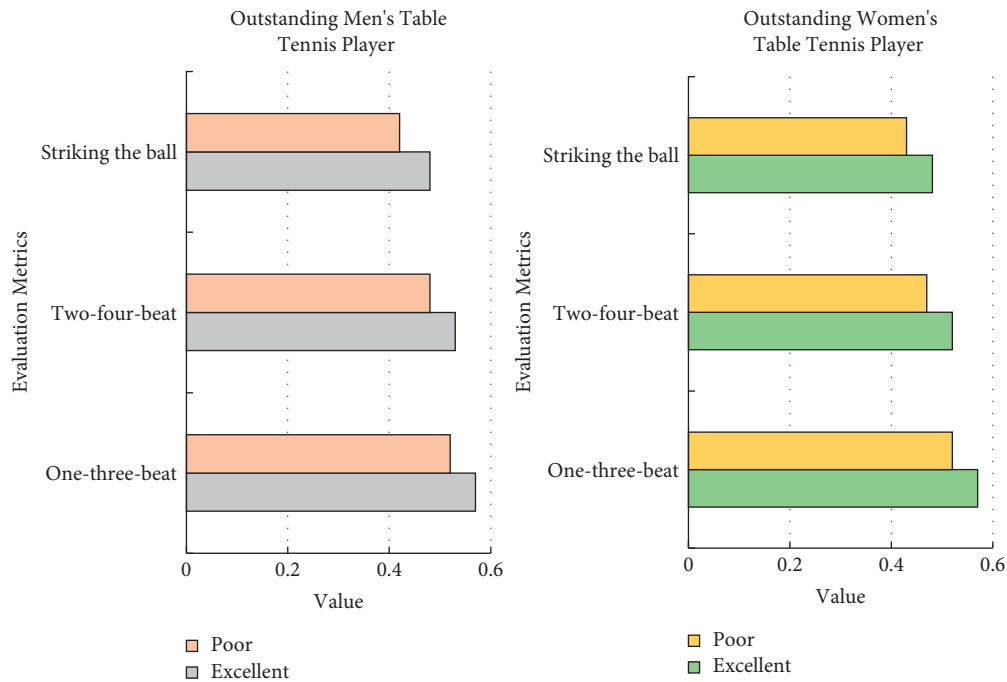
Figure 10: Evaluation criteria for TE value of technical performance of outstanding male and female table tennis players.

5. Conclusion

Table tennis is one of China's competitive sports advantages. As China's national team, table tennis has reached the peak of the international sports world again and again, leading the healthy and vigorous development of China's sports industry. This paper designs and implements a knowledge exchange platform for the training and competition of the Chinese table tennis team based on the current situation of information system applications in the field of sports and the needs of training and competition and team management of the Chinese table tennis team. Through the swarm intelligence algorithm, the game data, game videos, technical and tactical analysis data, training plans within the team, and other data are managed more efficiently and safely, which is a great saving in human and financial effort and helps the team train and play. In the system performance test, when the system load reached 50,000 users, it could still respond normally, fully meeting the performance requirements of the system. And it built a knowledge search engine and video search database based on table tennis, providing a novel knowledge acquisition channel and a personalized communication environment for coaches, athletes, scientific researchers, medical staff, and managers.

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] J. Zhang, "Automatic detection method of technical and tactical indicators for table tennis based on trajectory prediction using compensation fuzzy neural network," *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, Article ID 3155357, 12 pages, 2021.

[2] O. J. Choi, "Impact of the scoring system by stage on competition performance for national male table tennis player by country," *Korean Journal of Sports Science*, vol. 29, no. 5, pp. 1241–1249, 2020.

[3] H. Zheng, "Research on the importance and application of psychological factor training in table tennis teaching based on big data analysis," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 14, pp. 692–698, 2017.

[4] R. Yilmaz, F. G. Karaoglan Yilmaz, and E. Kilic Cakmak, "The impact of transactive memory system and interaction platform in collaborative knowledge construction on social presence and self-regulation," *Interactive Learning Environments*, vol. 25, no. 8, pp. 949–969, 2017.

[5] K. K. Bhattacharjee and S. P. Sarmah, "Modified swarm intelligence based techniques for the knapsack problem," *Applied Intelligence*, vol. 46, no. 1, pp. 158–179, 2017.

[6] H. Ma, S. Ye, D. Simon, and M. Fei, "Conceptual and numerical comparisons of swarm intelligence optimization algorithms," *Soft Computing*, vol. 21, no. 11, pp. 3081–3100, 2017.

[7] G. D. Ntouni, A. E. Paschos, V. M. Kapinas, G. K. Karagiannidis, and L. J. Hadjileontiadis, "Optimal detector design for molecular communication systems using an improved swarm intelligence algorithm," *Micro & Nano Letters*, vol. 13, no. 3, pp. 383–388, 2018.

[8] W. K. Mashwani, R. Haider, and S. Brahim Belhaouari, "A multiswarm intelligence algorithm for expensive bound

constrained optimization problems," *Complexity*, vol. 2021, no. 4, Article ID 5521951, 18 pages, 2021.

[9] B. Chen, H. Chen, and M. Li, "Improvement and optimization of feature selection algorithm in swarm intelligence algorithm based on complexity," *Complexity*, vol. 2021, no. 7, Article ID 9985185, 10 pages, 2021.

[10] Y. Gao, Y. Pan, H. Huang, E. R. Mohamed, and Z. M. Aly, "Swarm intelligence algorithm for extracting spatial spectrum features of hyperspectral remote sensing image and decomposing mixed pixels," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5045–5055, 2020.

[11] P. Gao, S. Wang, J. Lv, Y. Wang, and Y. Ma, "A database assisted protein structure prediction method via a swarm intelligence algorithm," *RSC Advances*, vol. 7, no. 63, pp. 39869–39876, 2017.

[12] X. Fu, M. Li, M. Yin, Q. Li, and Y. Chen, "IMRT images based on swarm intelligence algorithm in the treatment of nasopharyngeal carcinoma and pain," *Scientific Programming*, vol. 2021, no. 39, Article ID 7007169, 7 pages, 2021.

[13] P. A. Pasumpon, "Novel distance estimation based localization scheme for wireless sensor networks using modified swarm intelligence algorithm," *IRO Journal on Sustainable Wireless Systems*, vol. 2, no. 4, pp. 171–176, 2021.

[14] W. Zhong, D. Tan, X. Peng, Y. Tang, and W. He, "Fuzzy high-order hybrid clustering algorithm for swarm intelligence sets," *Neurocomputing*, vol. 314, no. 7, pp. 347–359, 2018.

[15] Z. Cao, Y. Xiao, M. Lu, X. Ren, and P. Zhang, "The impact of eye-closed and weighted multi-ball training on the improvement of the stroke effect of adolescent table tennis players," *Journal of Sports Science and Medicine*, vol. 19, no. 1, pp. 43–51, 2020.

Hindawi

Research Article

Segmentation and Classification of Encephalon Tumor by Applying Improved Fast and Robust FCM Algorithm with PSO-Based ELM Technique

Srikanta Kumar Mohapatra,[1] Premananda Sahu,[2] Jasem Almotiri,[3] Roobaea Alroobaea [ID],[3] Saeed Rubaiee [ID],[4] Abdullah Bin Mahfouz,[5] and A. P. Senthilkumar [ID][6]

[1]*Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India*
[2]*SRM Institute of Science and Technology, Ghaziabad, UP, India*
[3]*Department of Computer Science, College of Computers and Information Technology, Taif University, Taif 21974, Saudi Arabia*
[4]*Department of Industrial and Systems Engineering, College of Engineering, University of Jeddah, Jeddah, Saudi Arabia*
[5]*Department of Chemical Engineering, University of Jeddah, P.O. Box 80327, Jeddah 21589, Saudi Arabia*
[6]*Jigjiga University, Somali Regional State, East Africa, Jijiga, Ethiopia*

Correspondence should be addressed to A. P. Senthilkumar; senthilapsk@gmail.com

Received 30 May 2022; Revised 20 June 2022; Accepted 25 June 2022; Published 31 July 2022

Academic Editor: Amandeep Kaur

Nowadays, so many people are living in world. If so many people are living, then the diseases are also increasing day by day due to adulterated and chemical content food. The people may suffer either from a small disease such as cold and cough or from a big disease such as cancer. In this work, we have discussed on the encephalon tumor or cancer which is a big problem nowadays. If we will consider about the whole world, then there are deficiency of clinical experts or doctors as compared to the encephalon tumor affected person. So, here, we have used an automatic classification of tumor by the help of particle swarm optimization (PSO)-based extreme learning machine (ELM) technique with the segmentation process by the help of improved fast and robust fuzzy C mean (IFRFCM) algorithm and most commonly feature reduction method used gray level co-occurrence matrix (GLCM) that may helpful to the clinical experts. Here, we have used the BraTs ("Multimodal Brain Tumor Segmentation Challenge 2020") dataset for both the training and testing purpose. It has been monitored that our system has given better classification accuracy as an approximation of 99.47% which can be observed as a good outcome.

1. Introduction

Whenever we are going to do some investigation work, then first of all, we have to find out actually what is our motive. Here, our basic motivation is from magnetic resonance image (MRI), the pictures will segment first, the features or characteristics will remove, and finally that will be categorized whether it is cancerous or noncancerous.

To carry out the above issue, we have to use or apply some technique or tool which can all be corelated to each other. Initially, if we will discuss about the magnetic resonance image, then it is a leading technique to scan the whole brain and monitor the disease such as tumor. As our study is related to the neural network, it will explore the pictures and integrate them into 3-dimensional which can be used for better forecasting [1]. Next, the above processes have been used to help the clinical experts for detection of cancerous encephalon tumor in an effective manner. One more thing to be discussed here is that if the encephalon tumor spread steadily, some irregular activities have shown to the patients such as "vision problem and vomiting" which can cause a great problem if early investigation has not taken by the clinical experts [2]. If the physical examination has carried out from MRI, then it may be very heavy or time taking task

for the clinical practitioners due to composite shape of cancer area as well as some dotted spots inside the MRI picture. In this paper, some systematic work has been conducted that can determine that this is fully automatic and it may help the clinical experts. Initially, when somebody feels like drowsiness [3], memory problems, and some severe headache problem, there may be a chance for growing a tumor inside the brain [4–6]. So, that particular person has to visit the nearest hospital for scanning the brain. Regarding the scanning process, MRI is the best method to find out the clear picture of the brain as well as it has been built by some advance technologies [7]. For better understanding of MRI, we can use functional MRI due to the different types of activities that can be detected by changing with the blood vessels [8, 9]. After scanning, it needs to divide the pictures as some dotted or some odd pictures will present. For segmentation purpose, we have used the improved fast and robust fuzzy C mean technique as it can remove the noisy images [10]. After segmentation, some characteristics or features have to be removed from the input data so that the precision may be improved. For this purpose, we have used the frequently used technique gray level co-occurrence matrix. After removing some attributes, they will be nutrified as input to the particle swarm optimization based extreme learning machine for the categorization point of view which can be further divided into cancerous or noncancerous cells. This technique has been utilized to maximize the input and hidden layers that can be computed for tumor classification [11] as well as the bias along with weights. After successfully applied all the above techniques, it is the time for implementation process. Regarding this, we have used BraTs ("Brain Tumor Segmentation Challenges") for both the training as well as testing purpose.

The rest of this investigation work has been formulated in the following manner: Section 2 describes related work, Section 3 describes methodology, Section 4 describes result and discussion, and Section 5 describes conclusion followed by future work.

2. Related Work

Here, the authors mainly focused into segmentation and feature removal followed by categorization. Previously, some authors have worked in the above processes and we are going to discuss some of them as follows:

The authors in [12] have developed a compound forecasting model for perception and segmentation of brain tumor. In this work, they have used the hierarchical K-means clustering method with super rule tree for the above prediction process. Here, the FCM method has been used for the segmentation method which gives as an approximation of 88.9% of accuracy as well as they have concluded that their model can be used for multifaceted resection process also. Hua et al. [13] proposed the image segmentation method by a novel approach such as the improved multiview FCM clustering method. In their investigation work, they have also compared with the several clustering algorithms and their proposed model gives best accuracy. For implementation purpose, they have used the

Simulated Brain Database but for the high-dimensional multiview function, their proposed model may give some less accuracy. Hu et al. [14] has been investigated the encephalon disease forecasting process by the help of fuzzy system. Here, they have forecasted the image segregation regarding encephalon tumor process by "Hybrid Pyramid U-Net" model. They have also compared their proposed model with "local density clustering FCM and adaptive FCM" for their experimental purpose. Their model has given better classification accuracy but for the shape and texture characteristics, it has given some less accuracy.

Kumar et al. [11] has been proposed an automatic forecasting model for encephalon tumor identification and categorization. Here, they have examined a lot of things regarding the above process, but our main intention in this survey is to only the feature extraction process. In this work, they have used the mostly common method, gray level co-occurrence matrix, for feature reduction and it gives an awesome performance. They have again suggested that if some texture and shape features can be used in addition to GLCM, and then it may give best accuracy. Sathi and Islam [15] classified the encephalon tumor by ANN and extracted the features by some hybrid methods such as "Gabor filter and discrete wavelet based as well as GLCM" where they have used 6 characteristics by which the features has been extracted. Again, they have indicated that after the features have been extracted, it has given an awesome performance to detect the brain tumor. Mall et al. [16] extracted all the features from an X-ray image. Here, they have also used the GLCM technique for the above purpose. In their work, they have classified bone fracture instead of encephalon tumor detection. They have used 12 characteristics for the extraction purpose and have noticed a good result.

Velmurugan and Velmurugan [17] proposed the encephalon tumor classification process by the particle swarm intelligence method. By their investigation method, axial as well as coronal plane also detected. Their model has also given better accuracy rate, and some more machine learning approaches can be used for sagittal plane which can dissociate into sections from the left and right side of encephalon then it may give more accuracy rate. Khan et al. [18] has classified the encephalon cancer process by the extreme learning method with the association of correntropy via the mutual learning process. In their work, they have also used the BraTs dataset for experimental purpose and their outcomes are immovable. Ayane et al. [19] proposed the classification of encephalon tumor by the PSO-ELM technique and executed across the embedded system. They have used here the dataset "Harvard Medical School" for both the training as well as testing purpose. According to the researchers, their work has produced great achievement for finding the outcome and it may be helpful to the clinical experts for classifying the encephalon tumor as cancerous or noncancerous [20, 21].

From the above survey, it has been observed that for detection and classification of encephalon tumor not only the feature removal or classification are required but also some extra attributes such as texture characteristics and
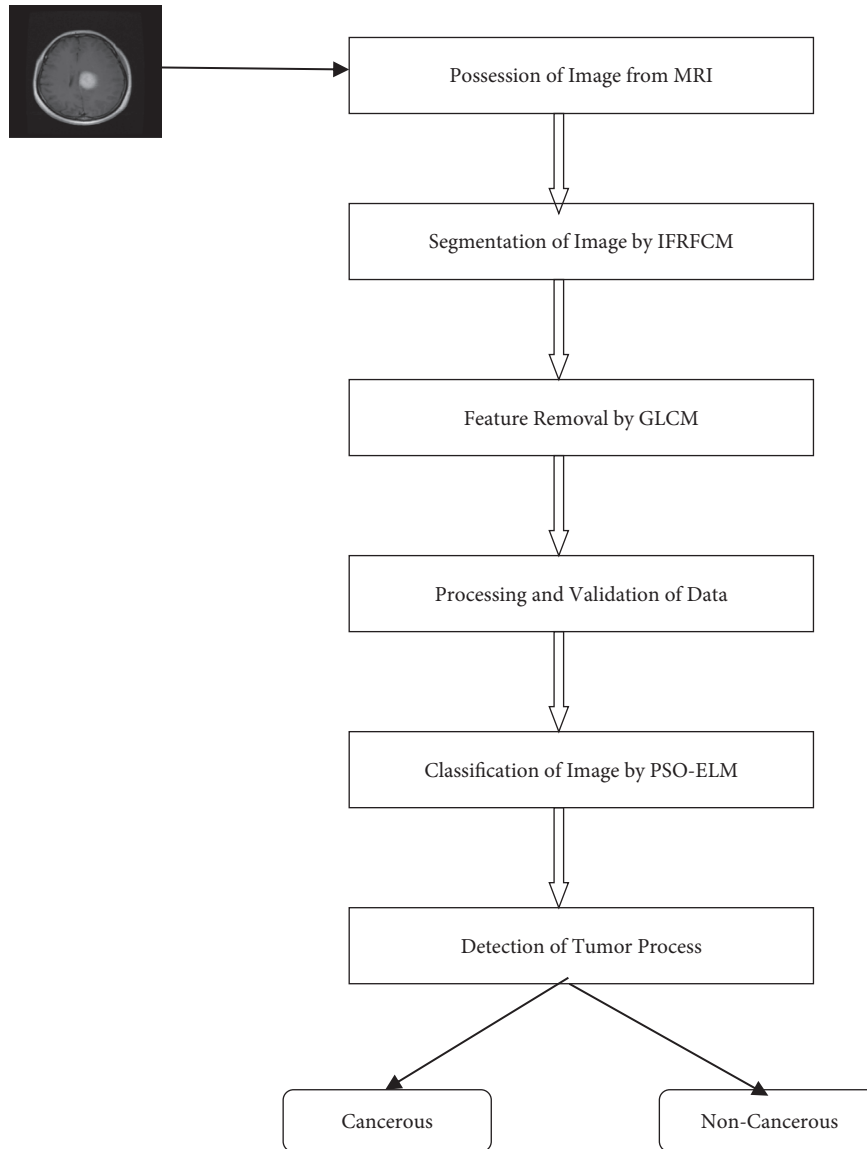
FIGURE 1: Workflow diagram of planned methodology.

coronal as well as sagittal suture are also used which we are going to implement in our work.

3. Methodology

In this work, mainly the researchers have focused into noncancerous tissues which may be found from segmentation [22] as well as the feature reduction method via magnetic resonance image [23]. This investigation work has flaunted in Figure 1.

This above diagram has been expanded through the following steps:

(i) Initially, the image has possessed by MRI [24].

(ii) Then, this image has been segmented by our IFRFCM technique [25].

(iii) Next, some features have removed by the most common method GLCM. [26].

(iv) After successful removal of features, the data then processed and proceeded for validation.

(v) The next process is classification which can be estimated by PSO-ELM. [27].

(vi) Finally, the tumor has been detected as cancerous or noncancerous.

3.1. Proposed IFRFCM for Image Segmentation. If the tumor is going to be detected from MRI, some noise has to be removed from those images. As described by the author [2], the above technique has the ability to remove the noise. As this technique supports "Wiener Filter" that has the ability to

observe the noisy images, here the objective function must be required. Now, the objective function as per FCM is [28, 29]

$$J_0 = \sum_{p=1}^{N} \sum_{q=1}^{C} \sigma_{qP}^v x_p - p_q^2 + \sum_{p=1}^{N} \sum_{q=1}^{C} G_{pq}, \quad (1)$$

where G_{pq} = fuzzy factor and it is defined as

$$G_{pq} = \sum_{\substack{r \in N_p \\ p! = r}} \frac{1}{d_{pr} + 1} \left(1 - \sigma_{qr}\right)^v x_r - p_q^2, \quad (2)$$

where

x_p, x_r = pixels

d_{pr} = spatial Euclidian distance between x_p and x_r

N_p = set of neighbors

σ_{qp} = Fuzzy membership of pth pixel with respect to q

σ_{qr} = neighbor of σ_{qp}

N = total number of pixels in image

C = cluster centre

V = fuzziness of significant division

With the diminish of analytical intricacy, the fuzzy division matrix is defined as

$$G'_{pq} = \sum_{\substack{r \in N_p \\ p! = r}} \frac{\log(\mu)^\alpha}{\exp(d_{pr}) + 1} \left(\sigma_{qr}\right)^v x_r - p_q^2, \quad (3)$$

where

μ = gray value of image

α = precision variable between 0 and 1

As we have to extract the marked objects from image, so some additional operations such as dilation and erosion are required which belongs to morphological reconstruction. Now, this type of image is indicated as γ_z and it is described as [30–32].

$$\gamma_z = R_d^c(f), \quad (4)$$

where

R_d^c = morphological closing reconstruction

f = original image

Now, by the help of morphological closing reconstruction, the above objective function may be adjusted as

$$J_0 = \sum_{q=1}^{C} \sum_{z=1}^{h} \left[\sigma_{qz}^v \gamma_z - p_q^2\right] + \sum_{q=1}^{C} \sum_{z=1}^{h} G_{qz}, \quad (5)$$

where σ_{qz} = degree of membership of gray value z in cluster q

Now, here, the fuzzy factor may be adjusted as

$$G'_{qz} = \sum_{\substack{r \in h_p \\ p! = r}} \frac{\log(\mu)^\alpha}{\exp(d_{pr}) + 1} \left(\sigma_{qr}\right)^v \gamma_z - p_q^2. \quad (6)$$

Now, the membership partition matrix appeared as S and it can be described as

$$S = \left[\sigma_{qr}\right]^{c*h}. \quad (7)$$

As we have discussed above, the Wiener Filter has the ability to observe the noisy image [20], so the new membership partition matrix S' can be described as

$$S' = \text{Wiener}[S]. \quad (8)$$

3.2. Proposed GLCM Technique for Feature Removal. As MRI consists of so many pictures that has to be captured, so for reduction of large number of pictures to the noisy picture, there is a removal technique is used [33, 34]. This is called as feature extraction process [35]. In this work, we have used the GLCM technique for feature extraction where it collects superior stage of facts of a picture such as structure, size, and color etcetera [36].

The previous researchers have used either only the GLCM technique or at max one of the feature like text characteristics or coronal or sagittal suture may be used [37]. However, here we have used all the above three features have been used for feature removal. All the process has described below:

(i) In texture features, initially, by the help of image processing technique, the images are combined to find the best outcome. Finally, the GLCM technique calculates the gray image and some texture features have been removed [38].

(ii) In coronal plane, the MRI images have been divided into front and back where the images are clearly visible and easy for removal [39].

(iii) In sagittal suture, the MRI images have divided into left from right so that the actual contusions are clearly visible [40].

From the above processes, it comes to know that GLCM with some characteristics have the ability to remove the features.

3.3. Proposed PSO-ELM Technique for Image Classification. Here, we have used one of the new techniques for classifying the tumor as cancerous or noncancerous, i.e., PSO-ELM. As PSO is an optimization technique, it has the ability to optimize the ELM parameters like weight and bias as well as the hidden neurons. From those hidden neurons, the above technique can find the good accuracy [41]. Here, the automatic categorization of encephalon tumor accomplished through our suggested PSO-ELM model [42]. Here, the weights are updated by the PSO technique [43] for finding each particle's best position by calculating the mean, so that it can upgrade to high competence.

It is a very useful as well as proficient training algorithm. Now, the output of ELM, Y, is determined as

$$Y = \sum_{i=1}^{H} v_i b_i(x) = \sum_{i=1}^{H} v_i b(w_i \times x_J + k_i), \quad J = 1, 2, 3, \ldots, N, \quad (9)$$

where H = number of hidden units, N = number of training samples, w = weight vector between input and hidden layer, v = weight vector between hidden and output layer, b = activation function, k = bias vector, and x = input vector.

Now, if we will assume beta matrix as special matrix, the output Y will be considered as

$$Y = Z * \beta, \quad (10)$$

where z = hidden layer output matrix and it can be described as

$$Z = \begin{bmatrix} b(w_1 \times x_1 + k_1) \cdots \\ \vdots \\ b(w_1 \times x_1 + k_1) \cdots \\ b(w_H \times x_1 + k_H) \\ \vdots \\ b(w_H \times x_N + k_H) \end{bmatrix}_{N \times H}. \quad (11)$$

Now, equation (11) is a linear system which may be solved by

$$\beta = z\dagger d, \quad (12)$$

where $z\dagger$ = Moore-Penrose Generalized Inverse of Matrix z and

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}. \quad (13)$$

3.3.1. Weight Optimization by PSO. As we know that particle swarm optimization is an optimization technique [44], which can optimize a problem by repeatedly testing to enhance a candidate solution with relation to a given measure of quality [45]. Basically, the design process is based on the manner of birds, so that they can exchange the information by swarming behavior [46, 47].

Previously, we have suggested that by the help of PSO, the best position [48] can be estimated by improvement of the weights. So, weights must be optimized in case of PSO. Here, each particle has factors such as position vector p_k as well as velocity vector w_k [49, 50] and these [51] are expressed in

$$w_k^{n+1} = w_k^n + \beta \delta_1 [g^* - p_k^n] + \gamma \delta_2 [p_k^* - p_k^n], \quad (14)$$

where δ_1 and δ_2 = random vectors between 0 and 1 [52].

g^* = global best position

β and γ = learning parameters

In a typical PSO algorithm, if an inertia function $\phi(n)$ is used, w_k^n has been replaced by $\phi(n) w_k^{n \cdot w}$ [53]. Now, this can be expressed as

$$w_k^{n+1} = \Phi w_k^n + \beta \delta_1 [g^* - p_k^n] + \gamma \delta_2 [p_k^* - p_k^n], \quad (15)$$

where $\Phi \in (0, 1)$.

Now, the weights are previously mentioned in Figure 2 as $w = [w_{i0} + w_{i1} x_1 + \cdots\cdots + w_{ij} x_j]$.

Furthermore, the weights are updated by using

$$w_i^{n+1} = (1 - \gamma) w_i^n + \gamma g^* + \beta \delta_n. \quad (16)$$

4. Result and Discussion

Here, we have used the BraTs 2020 dataset where it has used 278 MRI samples has been taken for both training and testing. In this work, 47 patient's data have been recorded in which 29 images for testing as well as 18 images for testing. All the exploratory analysis has been done with the python language with Google Co-lab in front end along with tensor flow tool in back end. Now, for the image segmentation by our planned method, IFRFCM has been presented in Figure 2.

Now for estimating the accomplishment of classification, we have taken one new technique that encompasses some additional things. It is purely based upon medical calculation and it is statistical software. Here, some new technical terms have been used and these are listed below:

For the diseased group, let

$$\text{Test Positive} = M,$$
$$\text{Test Negavtive} = N. \quad (17)$$

For the nondiseased group, let

$$\text{Test Positive} = P,$$
$$\text{Test Negavtive} = Q. \quad (18)$$

Disease prevalence: it is a statistical method which permits either the clinical experts or the researchers to estimate a person's chances of contracting a disease.

$$\text{Sensitivity} = \frac{M}{M + N},$$
$$\text{Specificity} = \frac{Q}{P + Q}. \quad (19)$$

Positive predicted value (PPV): if the test case is positive, the probability for that particular disease is present and it is described as

$$PPV = \frac{\text{Sensitivity} \times \text{Prevalence}}{\text{Sensitivity} \times \text{Prevalence} + (1 - \text{Specificity}) \times (1 - \text{Prevalence})}. \quad (20)$$

(a)   (b)   (c)
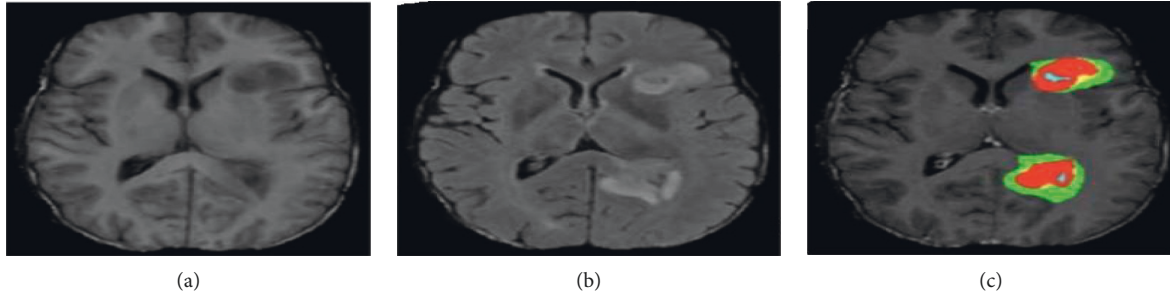
FIGURE 2: Segmentation outcome. (a) Original image. (b) Noisy image. (c) IFRFCM outcome.

TABLE 1: Estimation of various classification accomplishments.

Techniques	Sensitivity	Specificity	PPV	NPV	Accuracy
PSO	0.91	0.84	96.23	93.35	95.61
SCA	0.93	0.86	96.38	93.42	98.01
ASCA	0.94	0.88	96.47	9351	98.32
ELM	0.97	0.91	97.82	94.73	98.88
PSO-based ELM	0.99	1.0	98.35	95.69	99.47
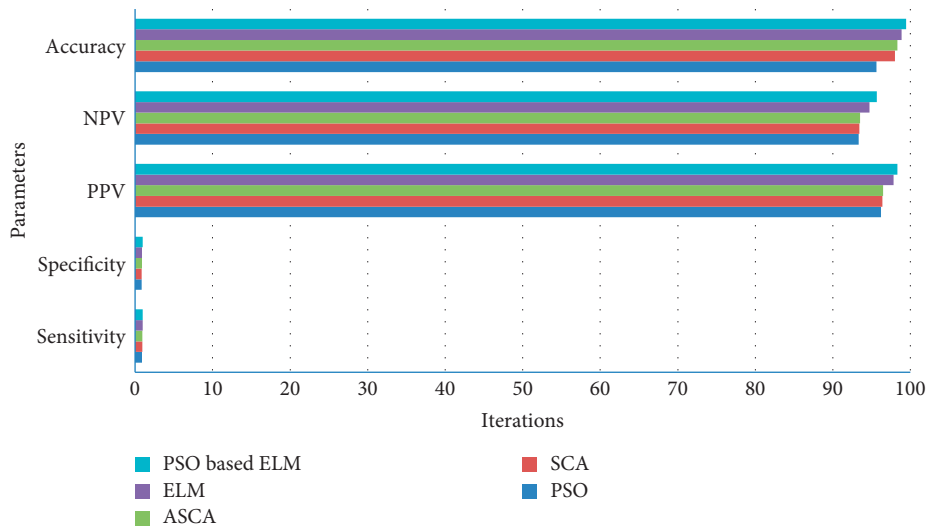
FIGURE 3: MSE of different techniques.

Negative predicted value (NPV): if the test case is negative, the probability for that particular disease is not present and it is described as

$$NPV = \frac{Specificity \times (1 - Prevalence)}{(1 - Sensitivity) \times Prevalence + Specificity \times (1 - Prevalence)}, \tag{21}$$

$$Accuracy = Sensitivity \times Prevalence + Specificity \times (1 - Prevalence). \tag{22}$$

By the above descriptions and from the BraTs dataset, the accomplishments of various classification methods are estimated in Table 1.

From Figure 2, it has been monitored that the IFRFCM segmentation technique has the ability to detect the noisy image as well as the clear cancer region [54]. By the help of

different parameters and number of iterations depicted in Figure 3, it has been observed that PSO-based ELM gives better classification accuracy as compared to the other classification models.

5. Conclusion

This research work has suggested 4 methods for the encephalon tumor detection process. Initially, the tumor region has been captured by MRI, all the images were divided, and the noisy images have been clearly found out by using the IFRFCM algorithm. After that, the features have been removed by the most commonly method, the GLCM method. Finally, the categorization or classification for finding out whether the tumor is cancerous or noncancerous has been investigated by the PSO-based ELM technique which gives an average classification accuracy rate of 99.47. Again, this proposed method can give better accomplishments if some deep characteristics have been applied.

Data Availability

The data will be available from the first author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] https://zhangruochi.com/Brain-Tumor-Auto-Segmentation-for-Magnetic-Resonance-Imaging-MRI/2020/04/17/#1.

[2] S. Mishra, P. Senapati, and M. R. Senapati, "MASCA-PSO based LLRBFNN model and improved fast and robust FCM algorithm for detection and classification of brain tumor from MR image," *Evolutionary Intelligence*, vol. 12, no. 4, pp. 647–663, 2019.

[3] A. Nowogrodzki, "The world's strongest MRI machines are pushing human imaging to new limits," *Nature*, vol. 563, no. 7729, pp. 24–26, 2018.

[4] L. Loued-Khenissi, O. Preuschoff, and K. Preuschoff, "An overview of functional magnetic resonance imaging techniques for organizational research," *Organizational Research Methods*, vol. 22, no. 1, pp. 17–45, 2019.

[5] T. Lei, X. Jia, Y. Zhang, L. He, H. Nandi, and A. K. Nandi, "Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 5, pp. 3027–3041, 2018.

[6] N. Varuna Shree and T. N. R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network," *Brain informatics*, vol. 5, no. 1, pp. 23–30, 2018.

[7] D. E. Ratnawati, M. Marjono, W. Widodo, and S. Anam, "PSO-ELM with time-varying inertia weight for classification of SMILES codes," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 6, pp. 522–532, 2020.

[8] S. Saraswathi, S. Sundaram, N. Sundararajan, M. Nilsen-Hamilton, and M. Nilsen-Hamilton, "ICGA-PSO-ELM approach for accurate multiclass cancer classification resulting in reduced gene sets in which genes encoding secreted proteins are highly represented," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 2, pp. 452–463, 2011.

[9] Y. Xue, C. Bai, D. Qiu, F. Li, and Z. Li, "Predicting rockburst with database using particle swarm optimization and extreme learning machine," *Tunnelling and Underground Space Technology*, vol. 98, Article ID 103287, 2020.

[10] https://www.kaggle.com/awsaf49/brats20-dataset-training-validation?select=BraTS2020_TrainingData.

[11] A. Kumar, A. Ashok, and M. A. Ansari, "Brain tumor classification using hybrid model of PSO and SVM classifier," in *Proceedings of the 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pp. 1022–1026, Greater Noida, India, October 2018.

[12] S. Simaiya, U. K. Lilhore, D. Prasad, and D. K. Verma, "MRI Brain Tumour Detection & Image Segmentation by Hybrid Hierarchical K-Means Clustering with FCM Based Machine Learning Model," *Annals of the Romanian Society for Cell Biology*, pp. 88–94, 2021.

[13] L. Hua, Y. Gu, X. Gu, J. Ni, and T. Ni, "A novel brain MRI image segmentation method using an improved multi-view fuzzy c-means clustering algorithm," *Frontiers in Neuroscience*, vol. 15, Article ID 662674, 2021.

[14] M. Hu, Y. Zhong, S. Xie, H. Lv, and Z. Lv, "Fuzzy system based medical image processing for brain disease prediction," *Frontiers in Neuroscience*, vol. 15, Article ID 714318, 2021.

[15] K. A. Sathi and M. S. Islam, "Hybrid feature extraction based brain tumor classification using an artificial neural network," in *Proceedings of the 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, pp. 155–160, Greater Noida, India, October 2020.

[16] P. K. Mall, P. K. Singh, and D. Yadav, "Glcm based feature extraction and medical x-ray image classification using machine learning techniques," in *Proceedings of the 2019 IEEE Conference on Information and Communication Technology IEEE*, pp. 1–6, Allahabad, India, December 2019.

[17] S. Velmurugan and T. Velmurugan, "Detection of brain tumor by particle swarm optimization using image segmentation," *Indian Journal of Science and Technology*, vol. 8, no. 22, 2015.

[18] M. A. Khan, I. Ashraf, M. Alhaisoni et al., "Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, p. 565, 2020.

[19] T. H. Ayane, "Brain tumor detection & classification using FRFCM segmentation and PSO based extreme machine learning and it's implementation through embedded system," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 13, pp. 452–465, 2021.

[20] D. C. Pereira, R. P. do Nascimento, and M. Z. Do Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 114, no. 1, pp. 88–101, 2014.

[21] S. A. Alazawi, N. M. Abbas, and A. H. Abbas, "Texture features extraction based on GLCM for face retrieval system,"

Periodicals of Engineering and Natural Sciences, vol. 7, no. 3, p. 1459, 2019.

[22] G. J. Djuričić, M. Radulovic, J. P. Sopta, M. Nikitović, and N. T. Milošević, "Fractal and gray level cooccurrence matrix computational analysis of primary osteosarcoma magnetic resonance images predicts the chemotherapy response," *Frontiers in Oncology*, vol. 7, p. 246, 2017.

[23] R. Maani, Y. H. Kalra, and S. Kalra, "Voxel-based texture analysis of the brain," *PLoS One*, vol. 10, no. 3, Article ID e0117759, 2015.

[24] A. H. Ali and M. Z. Abdullah, "An efficient model for data classification based on SVM grid parameter optimization and PSO feature weight selection," *International Journal of Integrated Engineering*, vol. 12, no. 1, pp. 1–12, 2020.

[25] N. Netjinda, T. Sirinaovakul, and B. Sirinaovakul, "Particle Swarm Optimization inspired by starling flock behavior," *Applied Soft Computing*, vol. 35, pp. 411–422, 2015.

[26] M. Safabakhsh, R. Safabakhsh, and A. Safabakhsh, "A novel stability-based adaptive inertia weight for particle swarm optimization," *Applied Soft Computing*, vol. 38, pp. 281–295, 2016.

[27] M. C. Software Ltd, "Diagnostic test evaluation calculator," p. 16, 2021, https://www.medcalc.org/calc/diagnostic_test.php.

[28] M. S. Mekala, G. Dhiman, G. Srivastava et al., "A DRL-based service offloading approach using DAG for edge computational orchestration," *IEEE Transactions on Computational Social Systems*, pp. 1–9, 2022.

[29] K. Yadav, A. Jain, N. M. Osman Sid Ahmed, S. A. Saad Hamad, G. Dhiman, and S. D. Alotaibi, "Internet of thing based koch fractal curve fractal antennas for wireless applications," *IETE Journal of Research*, pp. 1–10, 2022.

[30] S. Juneja, A. Juneja, G. Dhiman, S. Behl, and G. Kautish, "An approach for thoracic syndrome classification with convolutional neural networks," *Computational and Mathematical Methods in Medicine*, 2021.

[31] A. Juneja, S. Juneja, S. Kaur, and V. Kumar, "Predicting diabetes mellitus with machine learning techniques using multicriteria decision making," *International Journal of Information Retrieval Research*, vol. 11, no. 2, pp. 38–52.

[32] H. Upadhyay, S. Juneja, A. Juneja, G. Dhiman, and S. Kautish, "Evaluation of ergonomics-related disorders in online education using fuzzy AHP," *IETE Journal of Research*, 2021.

[33] S. Kanwal, J. Rashid, J. Kim, S. Juneja, G. Dhiman, and A. Hussain, "Mitigating the coexistence technique in wireless body area networks by using superframe interleaving," *IETE Journal of Research*, pp. 1–15, 2022.

[34] 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, pp. 1–10, Article ID 4912486, 2021.

[35] S. Juneja, S. Jain, A. Suneja et al., "Gender and age classification enabled blockschain security mechanism for assisting mobile application," *IETE Journal of Research*, pp. 1–13, 2021.

[36] S. Sharma, S. Gupta, D. Gupta et al., "Recognition of Gurmukhi Handwritten City Names Using Deep Learning and Cloud Computing," *Scientific Programming*, vol. 2022, Article ID 5945117, 2022.

[37] G. Dhiman, J. Rashid, J. Kim, S. Juneja, W. Viriyasitavat, and K. Gulati, "Privacy for healthcare data using the byzantine consensus method," *IETE Journal of Research*, pp. 1–12, 2022.

[38] H. Ding, X. Cao, Z. Wang et al., "Velocity clamping-assisted adaptive salp swarm algorithm: balance analysis and case studies," *Mathematical Biosciences and Engineering*, vol. 19, no. 8, pp. 7756–7804, 2022.

[39] G. Dhiman, S. Juneja, W. Viriyasitavat et al., "A novel machine-learning-based hybrid CNN model for tumor identification in medical image processing," *Sustainability*, vol. 14, no. 3, p. 1447, 2022.

[40] S. Sharma, S. Gupta, D. Gupta et al., "Deep Learning Model for the Automatic Classification of White Blood Cells," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7384131, 13 pages, Houssein, 2022.

[41] J. Rashid, S. Batool, J. Kim et al., "An augmented artificial intelligence approach for chronic diseases prediction," *Frontiers in Public Health*, vol. 10, Article ID 860396, 2022.

[42] Y. Alharbi, A. Alferaidi, K. Yadav, and S. Kautish, "Denial-of-Service Attack Detection over IPv6 Network Based on KNN Algorithm," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 8000869, 6 pages, 2021.

[43] S. Sharma, S. Gupta, D. Gupta et al., "Performance evaluation of the deep learning based convolutional neural network approach for the recognition of chest X-ray images," *Frontiers in Oncology*, vol. 29, 2022.

[44] H. Chugh, S. Gupta, M. Garg et al., "Image retrieval using different distance methods and color difference histogram descriptor for human healthcare," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–10, Article ID 9523009, 2022.

[45] V. K. Gupta, S. K. Shukla, and R. S. Rawat, "Crime tracking system and people's safety in India using machine learning approaches," *International Journal of Modern Research*, vol. 2, no. 1, pp. 1–7, 2022.

[46] T. Sharma, R. Nair, and S. Gomathi, "Breast cancer image classification using transfer learning and convolutional neural network," *International Journal of Modern Research*, vol. 2, no. 1, pp. 8–16, 2022.

[47] S. K. Shukla, V. K. Gupta, K. Joshi, A. Gupta, and M. K. Singh, "Self-aware execution environment model (SAE2) for the performance improvement of multicore systems," *International Journal of Modern Research*, vol. 2, no. 1, pp. 17–27, 2022.

[48] S. Ahmed, M. Monirujjaman KhanMonirujjaman Khan, R. Alroobaea, and M. Masud, "Development of a multi-feature web-based physiotherapy service system," *Intelligent Automation & Soft Computing*, vol. 29, no. 1, pp. 43–54, 2021.

[49] 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.

[50] A. Abbas, R. Alroobaea, M. Krichen, and R Saeed, "Blockchain-assisted Secured Data Management Framework for Health Information Analysis Based on Internet of Medical Things," *Pers Ubiquit Comput*, 2021.

[51] T. M. Ali, A. Nawaz, A. Ur Rehman et al., "A sequential machine learning-cum-attention mechanism for effective segmentation of brain tumor," *Frontiers in Oncology*, vol. 12, Article ID 873268, 2022.

[52] S. Safdar, M. Rizwan, T. R. Gadekallu et al., "Bio-Imaging-based machine learning algorithm for breast cancer detection," *Diagnostics*, vol. 12, no. 5, p. 1134, 2022.

[53] S. Abbas, Z. Jalil, A. R. Javed et al., "BCD-WERT: a novel approach for breast cancer detection using whale optimization based efficient features and extremely randomized tree algorithm," *PeerJ Computer Science*, vol. 7, Article ID e390, 2021.

[54] A. Dhankhar, S. Juneja, A. Juneja, and V. Bali, "Kernel parameter tuning to tweak the performance of classifiers for identification of heart diseases," *International Journal of E-Health and Medical Communications*, vol. 12, no. 4, pp. 1–16, 2021.

Research Article

Detection of Pneumonia Infection by Using Deep Learning on a Mobile Platform

Alhazmi Lamia [ID] [1] **and Alassery Fawaz** [ID] [2]

[1]*Department of Management Information System, College of Business Administration, Taif University, P.O Box 11099, Taif 21944, Saudi Arabia*
[2]*Department of Computer Engineering, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to Alassery Fawaz; falasser@tu.edu.sa

Pneumonia is a disease that spreads quickly and poses a serious risk to the health and well-being of its victims. An accurate biomedical diagnosis of pneumonia necessitates the use of various diagnostic tools and the evaluation of various clinical features, all of which are hindered by the lack of available experts and tools. According to the research presented here, a mobile app that uses deep learning techniques to classify whether or not a patient has pneumonia is being developed. It was hoped that a mobile application prototype for detecting pneumonia using neural networks would be developed as part of this study. The use of a high-level tool such as Create ML makes this process easier and eliminates issues such as how many layers a neural network has, initializing the model parameters, or which algorithms to use. The model can now be accessed by anyone, anywhere, via a mobile application. The dataset of more than 5,000 real images was used to train an image classification model using Create ML, a tool with a graphical interface, and there was no need for specialized knowledge.

1. Introduction

In health services in Iraq, respiratory infections represent between 50 and 70% of consultations and between 30 and 60% of hospitalizations [1]. Among adults suffering from pneumonia, it is estimated that between 22 and 42% require hospitalization and between 5 and 10% need an intensive care unit, and the lethality varies between 5 and 50% depending on the severity of the condition, which is higher in the elderly and immunosuppressed patients [2–4].

The diagnosis of pneumonia requires a review of chest radiography (CXR) by a highly qualified specialist, laboratory tests, vital signs, and clinical history, which makes its detection a difficult task. It normally presents as an area of increased opacity within the CXR. Even so, the identification of the diagnosis of pneumonia is complex due to other pulmonary conditions such as hemorrhages, lung cancer, postsurgical changes, pulmonary edema, atelectasis, or collapse. The comparison of CRX performed at different times and the relationship with the clinical history is essential for diagnosis.

Taking advantage of the large number of images generated by digital processing [3], some studies [5–9] focus on methods based on convolutional neural networks (CNN) to define whether a patient has pneumonia or not since they learn and select functions automatically. Other works [10–13] highlight the analysis of cracks using artificial neural networks, hidden Markov model, modeling of Gaussian mixtures (Gaussian Mixture Models, GMM), and algorithm K-NN (K-Nearest Neighbors). Although the models that currently exist have achieved performances that rival expert doctors [14–16], these proposals mostly remain in algorithms or platforms and do not develop useful tools. Due to advances in machine learning and greater power in smartphones, some authors leave open the possibility of incorporating these algorithms, which in turn can generate a great demand for this type of tool [4].

This work seeks to design a mobile application for the automatic detection of pneumonia, which may be appropriate in regions where health professionals do not arrive or as support when defining a diagnosis to reduce the mortality rate associated with Pneumonia [17–20].

In this article, we present a prototype application for the support of medical personnel in the diagnosis of pneumonia through radiographic images of the patient's chest through the use of neural networks.

2. Theory

In this work, we present an application prototype to support medical personnel in the diagnosis of pneumonia through radiographic images of the patient's chest through the use of neural networks.

2.1. Pneumonia. Pneumonia is a form of acute respiratory infection that impacts the lungs, which are composed of alveoli that are small sacs. When healthy people breathe, their lungs should be allowed to fill with air [21]. People who have pneumonia have alveoli that fill with pus and fluid, which makes breathing painful and limits the amount of oxygen that can be absorbed. Pneumonia is the single most important contributor to infant mortality around the world.

2.2. Diagnosis of Pneumonia. It is detected from an interrogation or physical examination, where the doctor listens to the lungs with the help of a stethoscope to identify crackling, bubbling, whistling, and rumbling sounds when inhaling and exhaling.

To make an accurate diagnosis, it is recommended to carry out some tests, which can be chest X-ray, pleural fluid culture, pulse oximetry, blood oxygen measurement, and bronchoscopy.

2.3. Machine Learning. Automatic learning (Machine Learning) seeks to create techniques or systems that learn automatically, and through patterns, it analyzes and treats information. Once the patterns are identified, they are able to predict behaviors and continuously improve.

A task in machine learning known as supervised learning involves inferring a function from training data that has been labeled. In order to solve a problem that involves this learning, the following steps need to be followed in the correct order: determine the type of training example, assemble a training set, determine the input features, determine the structure of the learned function and the corresponding learning algorithm, finish the design, and evaluate the accuracy of the learned function.

Finding a hidden structure in data that does not have labels is the goal of unsupervised learning, which occurs when there is neither an error signal nor a reward with which to evaluate a potential solution. Methods such as clustering and latent variable models (including the expectation maximization (EM) algorithm, method of moments, and

blind signal separation techniques) are examples of approaches to unsupervised learning. [22–25].

2.4. Artificial Neural Network (ANN). They are computer systems [26–29] that try to emulate some of the functions of living organisms. This means that they are made up of elements that mimic (in basic functions) the behavior and organization of the organism. The human brain can learn from experience and generalize from previous input to completely new input to predict an outcome. A neural network gains experience by analyzing data (Training) to determine behavior rules [30–32], based on which it can make predictions about new cases.

2.5. Deep Learning. It is an area within the field of machine learning that brings together a set of algorithms and techniques; with this method, learning is sought through examples, obtaining patterns from large amounts of data [20]. In this technique, the artificial network of neurons is composed of several layers of processing in a hierarchical manner, each one from a lower level of abstraction to a higher one (inspired by the biological behavior of the brain by the interconnections between neurons), and each one of them with a task. Specifically, various architectures organize the connections of different layers to determine the direction and propagation of data. Convolutional networks, in particular, have quickly become a methodology of choice for image analysis.

2.6. Convolutional Neural Network (CNN). It is a type of neural network specialized in the classification of images, videos, sounds, or text; its structure is similar to the visual cortex of a biological brain, so they are especially useful in locating patterns in images, recognizing objects, faces and scenes, and are widely used in the field of artificial division.

2.7. Review of the Literature. In 2018, Jakovljevic carried out a study on the tools to help diagnose pneumonia, where the focus is given on the analysis of the signals coming from the human body based on audio: cough, breathing, and speech, for the said signals. Starting from the common method, such as auscultation, the generated signal is processed with digital treatment techniques to be displayed, stored, and reproduced. This information must be processed to filter and eliminate noise and distortion so that it can be used as a training database. Among the studies, the analysis of crackles (discontinuous sounds of short duration) stands out using artificial neural networks, the Hidden Markov Model, Gaussian Mixture Modeling (GMM), and K-NN Algorithm (K-Nearest Neighbors), where it is concluded that neural networks and the K-NN algorithm give the best results in terms of monitoring respiratory sounds. It leaves open the possibility of introducing algorithms in increasingly powerful smartphones [33].

By 2019, Medrano Roldán using Apache Spark analyzed data as a framework in memory based on distributed processing and establishes how to implement a convolutional

neural network to automatically classify radiographs of patients with pneumonia as the database that is stored increases. The analyzes concluded that the database contains "x-rays of patients with pneumonia regardless of whether it is due to viruses or bacteria, and patients with a normal diagnosis where it is finally reduced to two classes, with and without pneumonia." She divided into 10 data sets, reduced the total number of images from 5,856 to 3,166 to avoid an imbalance, and normalized the images by training and evaluating the network with each data set [12, 13].

Also, in 2019, Samir Yadav and Shivajirao Jadhav [14] compared studies where images are used to diagnose diseases, where they highlighted the implementation of deep neural networks (DNN), especially convolutional neural networks (CNN), and achieved significant performance since 2012. They show that some investigations on the classification of medical images by CNN have managed to rival human experts. Where they give us an example CheXNet, a CNN with 121 layers trained with more than 100,000 frontal view chest X-rays, obtained better performance than the average performance of four radiologists. He further states that Kermany proposes a transfer learning system to classify images where the weighted average error is equal to the average of 6 human experts.

Abate (2019) et al. [15] mentioned significant achievements of CNNs in large amounts of data sets, but for cases where the data set is small, it sometimes fails if proper care is not taken. They propose a transfer learning approach using pretrained architectures to obtain the same result between small and large ensembles. They used five pretrained models, where they took the combination of them to form a large ensemble architecture and achieved promising results.

From the previous works, we can conclude that the approach that has been given by different researchers over time has been aimed at establishing algorithms that can select the characteristics automatically and evaluations to reduce false positives when classifying the diagnosis of pneumonia, but none has made an application on a mobile device as a support tool for the detection of pneumonia.

3. Methodology Implementation

The project is developed under the CRISP-DM methodology [16] used for development projects in the area of Data Analytics; it describes the common activities that must be carried out for the completion of a project of this nature; it consists of the next six phases.

3.1. Business Understanding.
In this phase, the business objectives and current needs are determined. For this, it is necessary to carry out a review of the current technologies in the diagnosis of pneumonia, the actors, time, and availability not only for medical personnel but for the general public; with this information, the development is proposed and limited to the project.

3.2. Understanding the Data.
The data collection process is carried out by searching different sources of information.

The data used for the creation of our model was obtained from http://www.cell.com/cell/fulltext/S0092-8674(18)30154-5. These contain images of chest radiographs in patients with pneumonia of different causes and healthy patients.

3.3. Preparation of Existing Data.
To carry out the training, it is necessary to have a data set organized and classified in these two categories; it is also important to have a group of data for training and another set for tests. It is not recommended to use training data to carry out tests, but it is recommended to use 80% of data for training and 20% for testing. At least 10 images are required for each category (NORMAL and PNEUMONIA). The more data there is, the better the accuracy of the model; the number of images in each category must be balanced as far as possible; 10 images cannot be used in the NORMAL category and 1,000 images in the PNEUMONIA category.

The images can be in any jpeg, png, or other formats (Figure 1); the images do not have to be the same size or have a specific size. However, it is better to use images of at least 299×299 pixels.

3.4. Modeling.
For modeling, Create ML is used. For several years, Apple has been venturing into the world of machine learning and also offering tools so that developers can easily integrate these features into applications. In this way, at WWDC, I present the Create ML, an application to build, train, and implement machine learning models directly from a Mac, without the need for Internet and advanced knowledge in machine learning [17–21].

Create ML (Figure 2) is a new way to train custom machine learning models that offers a simple and intuitive interface generating an extremely simplified experience; you can train models to perform tasks such as recognizing images, extracting meaning from text, or finding relationships between numerical values, all it requires is a sufficient amount of data for training.

Create ML Takes advantage of the machine learning infrastructure integrated into Apple products such as Photos and Siri, among other frameworks that come directly with the device; this makes some models smaller, more efficient, and require much less training time.

Currently, there are other very popular tools for creating custom models, such as Tensorflow and Caffe, but these require a large amount of code and do not have a friendly interface, and often require command of a programming language such as Python and some libraries. There are also other cloud-focused tools like Google Cloud AutoML and IBM AutoAI, but these require a constant Internet connection for data loading and training.

In the case of Google Cloud AutoML and Create ML, the difference between the accuracy of the models obtained using these two tools is not significant; both tools are good enough, and the decision on which to choose depends on personal preferences, availability of Internet, equipment, and storage space.
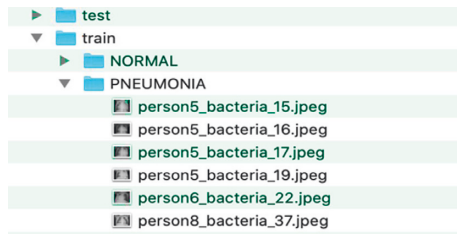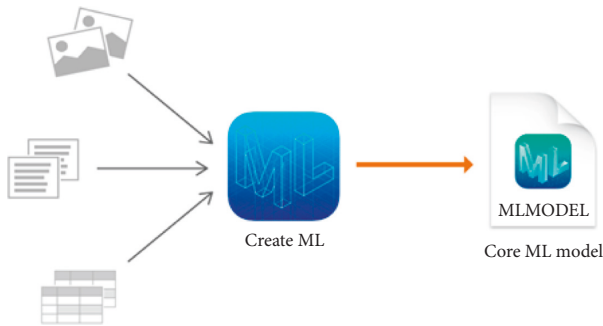
FIGURE 1: Data structure.

FIGURE 2: Creating ML from Apple (https://developer.apple.com/documentation/createml).

Create ML accepts various input types available for creating models images, sound, text, and number tables. In this case, a model will be used that uses an image as input and classifies it into two categories according to its content, NORMAL and PNEUMONIA.

3.4.1. Model Type Selection. To train a model using Create ML, it is necessary to create a project based on the different model templates; these are grouped according to the type of input of the model that we want to train.

After selecting "Image Classifier, "Select" Next", and we put some complementary data such as the name of the project, license, description of the model, and location of the project. Then the main interface is shown, which is quite simple and intuitive; you can train multiple versions of the same model using different datasets simultaneously. In the left panel, the different resources of the model are shown, and in the right panel, the data flow goes from left to right with the steps required to perform in the training, from inputting data to finishing with the output of the trained model (Figure 3(a)).

3.4.2. Entry. The input data for model training is divided into three groups:

(1) Training: classified data that the model will use to learn

(2) Validation: data that the model will use to verify the learning in each iteration, the application allows to take them randomly from the training set

(3) Tests: for the data that the application will use to verify the model's efficiency, it is important that none of these images are in the training set

To select the data, it is only necessary to drag and drop the folder with the images in the corresponding section.

Additionally, some options are allowed to be configured, such as the maximum number of iterations for the training and some filters that will be applied to each image to increase the amount of data when there is not a good number of data, and you want to improve model accuracy.

3.4.3. Training. Once the different data sets have been selected and the parameters configured, you can start training the model; you only need to press the "Training" button located on the toolbar (Figure 3(b)), and you can see the progress live as the model learns from the training images.

The model learning process is iterative; in each iteration, the model learns according to the characteristics of the training data set and verifies the accuracy with the validation data set because the validation data is generated randomly. The model may vary each time the model is trained.

The training time depends on factors such as the amount of data, number of iterations, filters applied to each image, and the team's capabilities.

3.4.4. Assessing Model Accuracy. Once the training is complete, you will be able to see different metrics of this process; creating ML shows the precision values with the training and validation data; this gives us information on how well the model is trained and allows us to make decisions.

The accuracy value for the training dataset is always close to 100% because the model has trained on these images. In our case, the model correctly identified 97% of the validation images and 86% of the test images.

The application also shows additional information on the training progress in each iteration, such as the number of elements in each category, precision, number of iterations, and total training time (Figure 4(a)).

Create ML to check the efficiency of the trained model against the test dataset, which contains images that the model has never seen before.

The model classifies all the images, and the app compares them to the correct category and presents the overall accuracy. For the first training, a general accuracy of 86% was obtained, and the accuracy for each category NORMAL (Healthy) 90% and PNEUMONIA 84% can also be observed; this indicates how good the model is to hit in each category (Figure 4(b)).

Suppose the performance of the model is very low, in that case, another training can be carried out with different parameters, increasing the number of iterations or with another more varied set of data, for example, images with different light exposures. This process can be carried out as many times as necessary to obtain the expected results. Create ML allows multiple training sessions to be carried out on the same model with different configurations and to compare the results of each of them in the same project. In this case, it can be seen that the performance can worsen depending on the parameterization (Figure 5(a)).
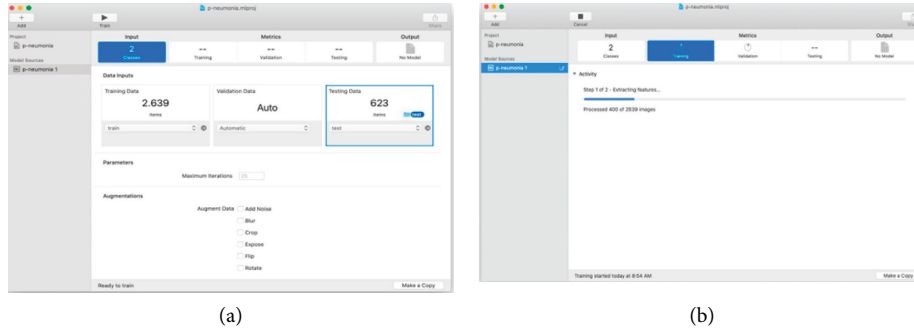
(a) (b)

FIGURE 3: (a) Configuring input data for model training and (b) model training.
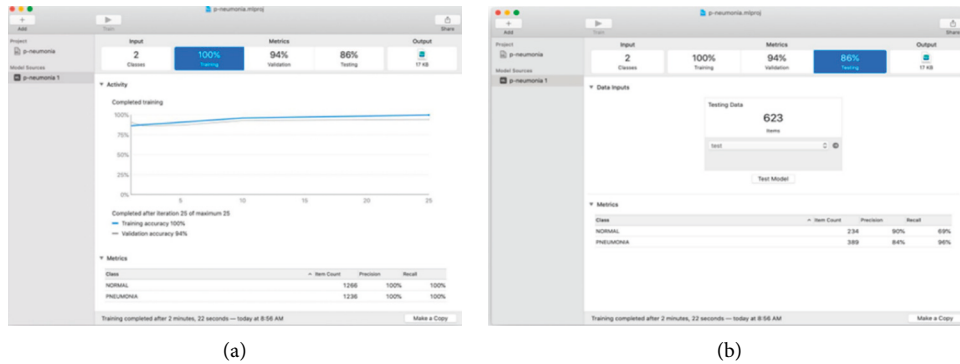

(a) (b)

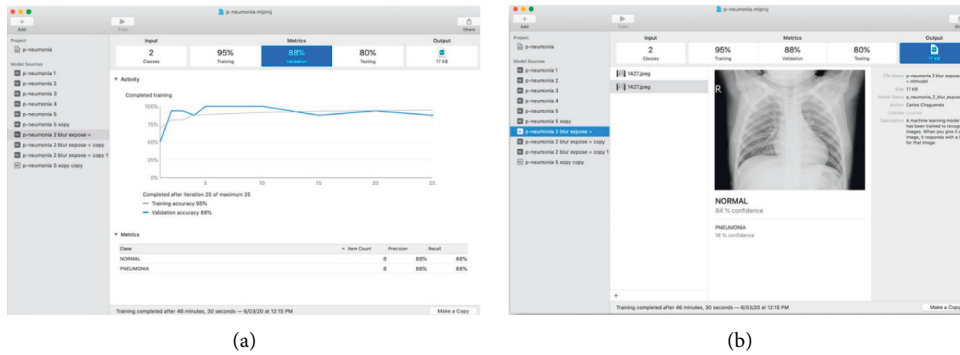FIGURE 4: (a) Training results; (b) model tests.

(a) (b)

FIGURE 5: (a) Retraining results with different parameters, (b) training result and tests with the model.

3.4.5. Exporting the Model. In the "output" section, the application allows testing with multiple images and shows the results of the classification (Figure 5(b)). When it is determined that the model works well enough, you can save the trained model to use; it is only necessary to drag and drop the model outside the application. It generates a file with the.mlmodel format, which is ready to be used in any application in the Apple ecosystem (iOS, iPadOS, macOS). One of the benefits of using Create ML to build models is how extremely small they are, and in this case, the model is only 17 KB.

3.5. Evaluation. Validation is performed according to expert judgment using the graphs obtained from the model accuracy training process to the group of test and validation images. At this point, it was possible to determine that the model is useful according to the needs of the business and meets the objectives to continue with the application prototype.

3.6. Implementation. Building the classification model by itself is not useful if it is not integrated into some application or service that uses it. This is why it is determined that a simple way to implement the model is through a mobile application; interested people can easily access it and use the device to capture or upload radiographic images for the model to make the prediction. The SCRUM methodology [34] is used to develop the application prototype.

Integrating pneumonia.mlmodel model in an iOS application is quite simple since the entire Apple ecosystem

FIGURE 6: Application operation in an iPhone 11 emulator.
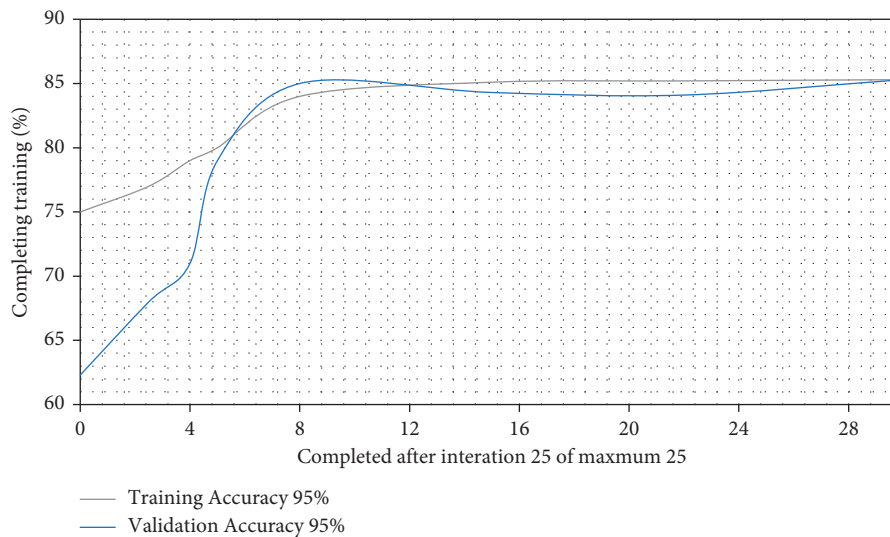

Training Accuracy 95%
Validation Accuracy 95%

FIGURE 7: Graph of training results of a model.

provides the tools to make this process as simple as possible; it can be used with a few lines of code.

Xcode is the integrated development environment offered by Apple for building applications. To use the model, it is only necessary to drag the model to a project in Xcode. Once the model is part of the project, Xcode displays the model information, type, description, size, input, and output data.

Thanks to the different frameworks included in Xcode, such as Core ML and Vision, classes are automatically created for the use of the model; other frameworks such as AVFoundation and PhotoKit allow us to obtain the images that are used as input for the model, whether they are from the device's camera or the photo library. Only a few lines are required to use the model in code, with which classification and a confidence percentage of this prediction are obtained from an image.

3.7. Functioning of the Application Prototype. The application consists of a few simple steps; the user can use the camera to take a photo or import a certain image from the photo gallery for analysis; once the user selects the image, the application processes it and displays the result with the prediction and the percentage of confidence offered by that result. The following are screenshots of the app on an iPhone 11 emulator, as shown in (Figure 6).

4. Results and Discussion

Data is the most important part of the project, and having a sufficient and reliable data set is the key to success. Performing the training of an image classifier model through the CreateML application is simple; unlike other tools, it abstracts the entire process in a simple interface, which means that great knowledge is not required to carry out the said process. The application allows the configuration of parameters and multiple workouts with different configurations; it generates graphs that allow us to make decisions, as shown in (Figure 7).

The classification models are composed of multiple layers in charge of carrying out simple processes and communicating these results to other layers to classify an image. The first layers are responsible for taking the raw pixel values and generating high-level abstract ideas such as "it's white" or "it's an animal" as you move between layers, and more specific details of the images are obtained until you can
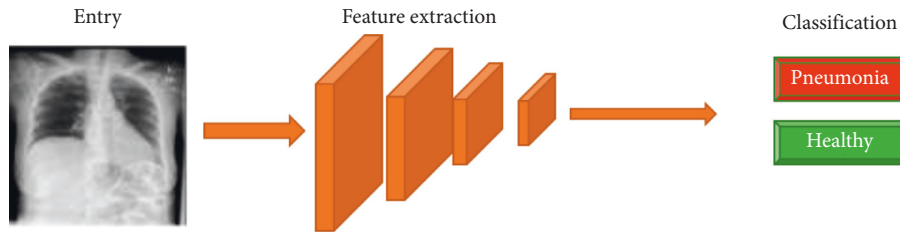
FIGURE 8: Classifier model training flow.

distinguish between NORMAL and PNEUMONIA (Figure 8). There are also different ways to organize and communicate the different layers and there are different architectures to organize these layers depending on the type of problem to be solved. Still, all this process is internal, thanks to CreateML.

After carrying out several training processes with different configurations and variations in the parameters, it was possible to obtain a model with an accuracy of ~85% integrated into an iPhone application that is very easy to use and distribute.

Since the model was trained with a set of data closed to two categories, NORMAL (Healthy) and PNEUMONIA, any image passed as input to the model will be categorized into one of these two groups since, during the training process.

If the image of a cat is sent, the model will respond incorrectly with the label PNEUMONIA. To solve this problem, you can choose to create another category of training data called OTHERS. It contains various images of things or objects that the user could pass as input to the model *to* improve the user experience.

It was also possible to show that factors such as lighting, contrast, blur, shooting angle, background, and different distortions present in the input image can affect the model's accuracy because within the training group, they are not being considered. Example images are those with these factors so that the model learns these characteristics with these variations. It is enough to include images of this type in the different data sets to solve this problem. The images present in the training process should be as close as possible to what is expected as input to the model, that is, how the end-user is expected to take the images.

5. Conclusions and Recommendations

5.1. Conclusions. The objective of this research work was to propose a mobile application prototype for the detection of pneumonia from a chest X-ray image through the use of neural networks. The model was created using Create ML, a high-level tool that simplifies this process and eliminates challenges, such as choosing how many layers a neural network has, initializing the model parameters, or what algorithms to use. The implementation of the model in a mobile application offers its use to anyone; the source code of the application is available on Github.

It was possible to establish through the review of the state of the art that there are various investigations to address the diagnosis of pneumonia using the field of artificial intelligence; some of them use radiographic image analysis and others stand out for analysing lung sounds. Many of these research projects focus on enhancing the efficiency of the models and end with their training. No research was found that uses the potential of mobile devices for the deployment of these models, although some authors leave this possibility open. This research offers an approach to the said processes to fill this existing gap.

Although artificial intelligence has been talked about since the middle of the last century, the most important advances have been made in recent years with the improvement of technique and technology; these improvements have intensified its use in different fields such as image recognition, voice, stock prediction, text generation, language translation, fraud prevention, autonomous driving, genetic analysis, and disease prognosis, being in the field of medicine one of the most interesting as it can help both medical staff and patients to make quick and accurate decisions, thus being able to save lives by offering early treatment. Using the existing data, it was possible to train several models with different configurations whose accuracies vary between ~78% and ~85%; we can conclude that the model's capabilities can be expanded if sufficient training data is available.

The model that obtained the best results was implemented in a mobile application prototype. The device's camera or photo gallery can be used for the application to process and generate a forecast. Its use is quite simple and generates a good user experience, the size of the application is only 4.8 MB, and the response time using the model is less than 1 second.

Future work can use a richer and more varied data set to classify different types of lung diseases. It is also possible to analyze other types of images such as MRIs or mammograms to predict whether a patient is likely to develop cancer.

5.2. Recommendations. For future projects related to image analysis and classification, we recommended the following:

(1) The data groups must be balanced in all the categories you want to classify

(2) They must have the most examples for each category

(3) The training data must fit the expected data as input to the model

(4) Different environmental conditions, such as lighting, can affect the model results, so it is important to have images in different conditions in the training data

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.

Acknowledgments

References

[1] F. Lami, H. Rashak, H. A. Khaleel et al., "Iraq experience in handling the COVID-19 pandemic: implications of public health challenges and lessons learned for future epidemic preparedness planning," *Journal of Public Health*, vol. 43, no. Supplement_3, pp. iii19–iii28, 2021.

[2] U. Thiem, H. Heppner, and L. Pientka, "Elderly patients with community-acquired pneumonia," *Drugs & Aging*, vol. 28, no. 7, pp. 519–537, 2011.

[3] Y. Gupta and S. Agrawal, "A study of lung disease using image processing in big data environment," *IOP Conference Series: Materials Science and Engineering*, vol. 1022, Article ID 012030, 2021.

[4] V. Sirish Kaushik, A. Nayyar, G. Kataria, and R. Jain, "Pneumonia detection using convolutional neural networks (CNNs)," *Lecture Notes in Networks and Systems*, pp. 471–483, 2020.

[5] M. B. Alazzam and F. Alassery, "The dynamic movement of disaster management systems based on vehicle networks and applied on the healthcare system," *Applied Bionics and Biomechanics*, vol. 2021, Article ID 5710294, 8 pages, 2021.

[6] I. Cardoso, E. Almeida, H. Allende-Cid et al., "Analysis of machine learning algorithms for diagnosis of diffuse lung diseases," *Methods of Information in Medicine*, vol. 57, no. 05/06, pp. 272–279, 2018.

[7] U. Sait, S. Shivakumar, G lal kv et al., "A mobile application for early diagnosis of pneumonia in the rural context," in *Proceedings of the 2019 IEEE Global Humanitarian Technology Conference (GHTC))*, pp. 1–5, Seattle, WA, USA, October 2019.

[8] H. Al-kuraishy, T. Al-Maiahy, R. A. Musa, and Z. H. Ali, "COVID-19 pneumonia in an Iraqi pregnant woman with preterm delivery," *Asian Pacific Journal of Reproduction*, vol. 9, no. 3, p. 156, 2020.

[9] R. E. Al Mamlook, S. Chen, and H. F. Bzizi, "Investigation of the performance of machine learning classifiers for pneumonia detection in chest X-ray images," in *Proceedings of the 2020 IEEE International Conference on Electro Information Technology (EIT)*, Chicago, IL, USA, August 2020.

[10] F. Shi, J. Wang, J. Shi et al., "Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19," *IEEE reviews in biomedical engineering*, vol. 14, pp. 4–15, 2020.

[11] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, M. Mahmood Khalil, and V. Ramesh, "Zelalem meraf, vishal kumar, "dynamic systems enhanced by electronic circuits on 7D," *Advances in Materials Science and Engineering*, vol. 2021, Article ID 8148772, 11 pages, 2021.

[12] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajjej, and A. A. Shihab, "Applying dynamic systems to social media by using controlling stability," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4569879, 7 pages, 2022.

[13] W. Enbeyle, W. Enbeyle, A. S Al-Obeidi et al., "Trend analysis and prediction on water consumption in southwestern Ethiopia," *Journal of Nanomaterials*, vol. 2022, Article ID 3294954, 2022.

[14] A. S. Al-Obeidi and S. F. Azzawi, "A novel six-dimensional hyperchaotic system with self-excited attractors and its chaos synchronisation," *International Journal of Computing Science and Mathematics*, vol. 15, no. 1, pp. 72–84, 2022.

[15] K. R. Swetha, M. Niranjanamurthy, M. P. Amulya, and Y. M. Manu, "Prediction of pneumonia using big data, deep learning and machine learning techniques," in *Proceedings of the 2021 6th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatre, India, July 2021.

[16] T. B. Chandra and K. Verma, "Pneumonia detection on chest X-ray using machine learning paradigm," in *Proceedings of the 3rd International Conference on Computer Vision and Image Processing*, pp. 21–33, Beijing, China, May 2020.

[17] N. Jakovljević and T. Lončar-Turukalo, "Hidden Markov model based respiratory sound classification," in *Book: Precision Medicine Powered by pHealth and Connected Health*, pp. 39–43, Springer, Singapore, 2018.

[18] A. A. A. Rahim Al-Baldawi, "The possibility of implementing industrial incubators and their role in the development of small industry and medium in Iraq," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–22, 2022.

[19] B. A. M. Muhammad, "The role of universities in developing societies by accreditation on scientific research," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–19, 2022.

[20] M. B. Alazzam, H. Al Khatib, W. T. Mohammad, and F. Alassery, "E-health system characteristics, medical performance, and healthcare quality at Jordan's health centers," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–7, Article ID 5887911, 2021.

[21] W. O'Quinn, R. J. Haddad, and D. L. Moore, "Pneumonia radiograph diagnosis utilizing deep learning network," in *Proceedings of the 2019 IEEE 2nd international conference on electronic information and communication technology (ICE-ICT)*, pp. 763–767, IEEE, Harbin, China, January 2019.

[22] P. Mayorga, C. Druzgalski, R. Morelos, O. Gonzalez, and J. Vidales, "Acoustics based assessment of respiratory diseases using GMM classification," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6312–6316, IEEE Engineering in Medicine and Biology Society. Conference, Buenos Aires, Argentina, November 2010.

[23] K. Medrano Roldán, "Red neuronal convolucional en un ambiente pseudo-distribuido para la clasificación de radiografías de pacientes con neumonía," *Licenciatura en Ing. en Sist. Comput.*, vol. 7, 2019.

[24] M. Rasmi, M. B. Alazzam, M. K. Alsmadi, I. A. Almarashdeh, R. A. Alkhasawneh, and S. Alsmadi, "Healthcare professionals' acceptance Electronic Health Records system: c," *International Journal of Healthcare Management*, vol. 13, no. sup1, pp. 48–60, 2020.

[25] S. Yadav and S. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data*, vol. 6, no. 1, 2019.

[26] S. Rajaraman, S. Candemir, G. Thoma, and S. Antani, "Visualizing and explaining deep learning predictions for pneumonia detection in pediatric chest radiographs," in *Medical Imaging 2019: Computer-Aided Diagnosis*, vol. 10950, pp. 200–211, SPIE, Bellingham, March 2019.

[27] F. Zhang, "Application of machine learning in CT images and X-rays of COVID-19 pneumonia," *Medicine*, vol. 100, no. 36, 2021.

[28] A. Abate, P. Barra, S. Barra, C. Molinari, M. Nappi, and F. Narducci, "Clustering facial attributes: narrowing the path from soft to hard biometrics," *IEEE Access*, vol. 8, pp. 9037–9045, 2020.

[29] A. Mamra, A. S. Sibghatullah, G. P. Ananta, M. B. Alazzam, Y. H. Ahmed, and M. Doheir, "Theories and factors applied in investigating the user acceptance towards personal health records: review study," *International Journal of Healthcare Management*, vol. 10, no. 2, pp. 89–96, 2017.

[30] Ibm, *Manual CRISP-DM de Ibm Spss Modeler*, Vol. 56, IBM Corp, Armonk, 2012.

[31] E. Pérez-Rodríguez, L. Curto, F. Arias, A. Ladron de Guevara, and A. Bendito, "Noninvasive evaluation of mediastinal metastasés in bronchogenic carcinoma with 67Ga scanning," *Chest*, vol. 90, 1986.

[32] H. M. Ariza, V. R. Mozo, and H. M. Quintero, "Methodology for the Agile development of software based on a guide for the body of knowledge of scrum (SBOKTM Guide)," *International Journal of Applied Engineering Research*, vol. 13, no. 14, pp. 11479–11483, 2018.

[33] Ishaq Yousef Al-Ayash Abdul Rasoul Saad, "Social intelligence and its relationship to decision quality," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–22, 2022.

[34] M. Fadhil Jwaid, "An efficient technique for image forgery detection using local binary pattern (hessian and center symmetric) and transformation method," *Scientific Journal Al-Imam University College*, vol. 1, pp. 1–11, 2022.

Research Article

Design and Implementation of Personalized Push Service Based on Feature Extraction and Pattern Recognition

Moxuan Zhang (ID)

Department of Mechanical Engineering, Tianjin Renai College, Tianjin, China

Correspondence should be addressed to Moxuan Zhang; zhangmoxuan0420_@tju.edu.cn

Personalized push service is one of the more popular research and application fields, which has received more and more attention. Its application prospects are also more and more extensive. This research mainly designs and implements personalized push services through feature extraction and pattern recognition. In this study, the Chinese texts of user-visited pages are classified according to keywords, so as to obtain the user's interest characteristic data. Then, according to the frequency of each feature category, the weight of the user's interest feature is calculated, and the user's interest field is predicted and identified. After that, resources that match the user's interest field are pushed to it. In order to verify the effectiveness of the improved model, this study carried out experiments and comparisons on the precision rate, recall rate, and comprehensive classification rate of the original model and the improved model on the implemented personalized push service system. In the research, the error between the interest results under each interest topic in the test set and the results obtained by the statistical analysis of the training set is within a reasonable range, the maximum of which is about 5%. The accuracy of interest degree prediction in different scenarios can reach more than 90%, which directly confirms the good applicability and effectiveness of the analysis and calculation method and the constructed model for user interest in this study. The personalized push service framework proposed in this study has good application value in the field of time-sensitive information services.

1. Introduction

The personalized push service can analyze and extract the user's interest from the collected user-related information in order to further understand the user. In addition, user behavior analysis can explore the behavior characteristics and information needs of users. In this way, resource information that matches user interests can be discovered, and it can be actively pushed to users to meet their differentiated needs of users. Personalized push service is the concrete embodiment of information personalized service. With the rapid development of computer technology, computers with large memory and fast and accurate computing have appeared in our real life, which makes people's ways of obtaining information resources develop towards diversification. The user's basic information, interests, usage behavior, and other information are continuously collected by the computer, and different databases are established for different users, thereby truly providing different personalized push services for different users. With the development of network technology, users can also have different personalized page customization when using the personalized push system. Users can also choose different types of services such as resources they need according to their actual situation.

According to the different needs of different users, it is very necessary to provide personalized push services, change passive services into active services, and assist users to filter out accurate and useful information from a large amount of data. Wohllebe et al. believed that push notifications are the core function of mobile applications [1]. Lyu et al. proposed a new model called deep matching to ranking (DMR) [2]. Xiao-Long et al. proposed a task-oriented multicollaborative filtering algorithm [3]. Wang et al. proposed a collaborative filtering algorithm recommendation system based on device-edge-cloud federated learning [4]. Ting proposed two data dissemination models: a data pull model in which

mobile users acquire data from data providers [5]. Their research precision on personalized push services is not very high. Therefore, this paper reviews relevant materials and decides to use feature extraction pattern recognition to optimize personalized push services.

Most of the current systems obtain their personal information through the way of user submission and lack the ability to track user behavior and actively analyze their behavior and extract user features to extract pattern recognition, so the personalized information inside the system is static. Jenke et al. believed that emotion recognition can assess the user's state [6]. Borges et al. proposed a method aimed at extracting features to obtain disturbance information related to power quality [7]. Matthew et al. used micro-Doppler signals collected by multistatic radar to detect and distinguish hovering and flying micro-UAVs carrying different payloads [8]. Pei et al. proposed a SAR ATR feature extraction method based on two-dimensional principal component analysis [9]. In order to reduce the cost of labeling samples, Liu et al. proposed a new semisupervised algorithm (SNC) with neighborhood constraints [10]. Using feature extraction pattern recognition, it is necessary to establish feature constraints based on user attributes and user interest-based feature constraints. The feature extraction pattern recognition will be further explored in the following sections.

The identification of information features mainly requires users to actively submit real data that can reflect their preferences to the system. The data provided by the user mainly includes the user's personal attribute information, as well as the subjective display score and text evaluation made by the user. This paper analyzes and constructs the push service architecture. It also designs each module of the architecture separately, including five modules: log file preprocessing, user behavior feature analysis, literature resource feature analysis, information push, and real-time data update. Instances are verified for the personalized push service architecture. That is, taking a user as an example and analyzing the user's behavior log, interest features are extracted. Then, information resources consistent with their interest features are pushed. This proves the rationality and feasibility of the architecture. Both push service and web page advertisement push need to analyze user behavior and extract users' interest characteristics, so as to predict their areas of interest and push relevant information resources to them. When the number of neighbors is greater than 28, the recommendation accuracy of the feature extraction pattern recognition model in this study is the highest. When the number of neighbors is less than 28, in terms of recommendation accuracy, the recommendation accuracy of the feature extraction pattern recognition model proposed in this study is better than the recommendation accuracy of the K-means algorithm and Pearson correlation.

2. Design and Implementation of Personalized Push Service

2.1. Personalized Push Service.
Personalized push service actively sends information resources matching their interest

characteristic data to users in a purposeful and timely manner according to user needs. A personalized push service is suitable for all kinds of users, and it can push continuously updated information resources to users. The personalized push service has the following characteristics:

(1) *Differences According to Users.* The personalized push service is user-centric. On the basis of fully mining and researching the user's browsing habits, search, borrowing records, and other information, according to the user's personal needs, it will discover the resource information that matches the user's information needs and actively push the information to users in order to meet the needs of different users [11].

(2) *Initiative.* Initiative is the most fundamental characteristic of personalized push service. Personalized push service has changed the previous mode of passively providing information to users. Instead, the digital library actively pushes information resources to users. When new users or new resources appear, or when user behavior logs are updated, the library will analyze the characteristic attributes of relevant information resources for users to consult anytime, anywhere.

(3) *Efficiency.* Through personalized push service, users can obtain the information resources they need from massive information resources, which helps to improve the efficiency of user resource acquisition. In addition, when the information resources are updated, the resources will be analyzed and actively pushed to the users in need, thus avoiding the waste of user time.

Digital resource integration is an important aspect of personalized push services. Resource integration can integrate not only various types of resources, cluster, or classify resources but also related data resources together, so as to provide users with the appropriate information, shortening the time for users to search for information resources, which provides users with efficient and personalized services. Through the matching of users and resources, the most interesting information can be recommended to user groups with similar interests [12].

2.2. Feature Extraction Pattern Recognition.
The foundation of the user's interest model is the extraction of data features, and the recommendation system uses these data features as the basis for extracting user interests. The key to building a model is the representation method and structural feature information of the model, which directly relate to the availability and computability of the model. This research adopts the method of feature extraction pattern recognition to develop personalized push service and also constructs the key points of user portrait and personalized extraction.

The simplest Boolean expression is equality, which tests whether one value is the same as another. The Boolean weight (binary) is based on whether the feature word appears

in the text as the basis for weight assignment, and its formula is expressed as

$$W_{ij} = \begin{cases} 1, & TF_{ij} > 0 \\ 0, & TF_{ij} \le 0 \end{cases}, \qquad (1)$$

where W_{ij} represents the weight of feature word i in text j, and TF_{ij} represents the number of times feature word i appears in text j [13].

TF/IDF is used to evaluate the importance of a word to a text in the training set. Among them, TF represents the frequency of a certain word in the text, that is, the word frequency. The idea of the TF weight calculation method is that the frequency of feature words appearing in the text is proportional to its importance. That is, the higher the frequency of occurrence is, the more important the feature word will be in the text, and the greater its weight will be. Its calculation formula is as follows:

$$W_{ij} = TF_{ij}. \qquad (2)$$

The calculation method of the membership degree of the text to be classified in each category is as follows:

$$\text{score}(d_i, C_j) = \sum_{i,j=1}^{k} \text{sim}(d_i, d_j) y(d_i, C_j), \qquad (3)$$

where $y(d_i, C_j)$ is the classification value of the text pair category in the training set, and its value range is {0, 1} [14].

Given the mother wavelet function $\psi(t)$, on the continuous (a, b), the wavelet transform basis function can be defined in the time-frequency space as

$$\Psi_{ab}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \leftrightarrow \phi_{ab}(\Omega) = \sqrt{a}\phi(a\Omega)e^{-jb\Omega}. \qquad (4)$$

The inverse transform of the wavelet transform of the signal can be defined as

$$f(t) = \frac{1}{C_\Psi} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \frac{da\,db}{a^2} W^f a_0(a,b)\Psi_{ab}(t). \qquad (5)$$

The scale function also satisfies the orthogonal condition when the scale parameters are equal [15].

$$\int \phi_{mn}(t)\phi_{ms}(t)dt = \delta_{n-s}. \qquad (6)$$

The set of scaling functions can be defined as [16]

$$\{\phi_{mn}(t)\} = 2^{-(m/2)}\phi(2^{-m}t - n). \qquad (7)$$

In CRP, the local compactness of a sample is characterized by the following minimization problem [17]:

$$\min_{P} \sum_{i=1}^{n} P^T x_i - \sum_{j=1}^{n} s_{j,i} P^T x_j = tr(P^T S_L P). \qquad (8)$$

On the other hand, the maximization of the sample population divergence can be expressed as [18]

$$\max_{P} \sum_{i=1}^{n} P^T x_i - P^T m_2^2 = tr(P^T S_t P), \qquad (9)$$

where m is the population mean of all samples [19].

The formula for the predicted score is as follows [20]:

$$P_{u,k} = \overline{R_k} + \frac{\sum_{i \in I} \text{sim}(k,i)(R_{u,i} - \overline{R_i})}{\sum_{i \in I}(|\text{sim}(k,i)|)}, \qquad (10)$$

where $P_{u,k}$ represents the predicted score of user U for item k.

The calculation formula of the weighted slope one algorithm is as follows:

$$P_w(u)_j = \frac{\sum_{i \in R_j}(\text{dev}_{j,i} + u_j)\text{card}(S_{j,i})}{\sum_{i \in R_j}\text{card}(S_{j,i})}, \qquad (11)$$

where $P_w(u)_j$ represents the predicted rating of item j by target user u.

According to the Euclidean distance, the attribute similarity of i and j can be calculated, and the calculation formula is as follows:

$$\text{sim}_i(i,j) = \frac{1}{1 + d(i,j)}. \qquad (12)$$

Through the attribute similarity between items, the weighted sum method is used to calculate the user's predicted score for the item. The weighted average is to multiply each value by the corresponding weight, then add and sum to get the overall value, and finally divide by the total number of units. The specific calculation formula is as follows:

$$u_i = \frac{\sum_{j=1}^{n} u_j * \text{sim}_s(i,j)}{\sum_{j=1}^{n} \text{sim}_s(i,j)}. \qquad (13)$$

Manhattan distance represents the nonlinear distance between two data objects. The formula is as follows:

$$d(x_i, x_j) = \sum_{k=1}^{m} |x_{ik} - x_{jk}|. \qquad (14)$$

The variance weighted distance is as follows:

$$d(x_i, x_j) = \left[\sum_{k=1}^{m} \frac{(x_{ik} - x_{jk})^2}{s_k^2}\right]^{1/2}. \qquad (15)$$

Euclidean distance formula is as follows:

$$d(x_i, x_j) = \sum_{k=1}^{m} \frac{(x_{ik} - x_{jk})^2}{s_k^2}. \qquad (16)$$

The information entropy formula of information source X is as follows:

$$H(X) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i}, \qquad (17)$$

where n indicates that there are n categories in the sample set X, and $p(ai)$ indicates the frequency of the i-th category in

the sample set X. The formula of the weighted information entropy of user U_m and user U_n is as follows:

$$\text{JWDE}(U_m, U_n) = -\frac{2n}{N1 + N2} \sum_{i=1}^{N} p(d_i) \log_2 (p(d_i)) \times |d_i|. \tag{18}$$

According to the similarity between user m and its nearest neighbors, the classical nearest neighbor prediction method is used to predict the score of m on unknown items as follows:

$$P_{m,J} = \overline{R_m} + \frac{\sum_{i \in \text{NBS}} \text{Sim}(m, i)(R_{I,J} - \overline{R_I})}{\sum_{i \in \text{NBS}} |\text{Sim}(m, i)|}. \tag{19}$$

2.3. Personalized Push Service Architecture. The architectural design of the personalized push service is shown in Figure 1. First, the behavior data (log files) generated is preprocessed during the user's use. Then, the content of the resource page accessed by the user is captured by the web crawling technology, and the Chinese text word segmentation is performed on the page. Then, the weight of each feature word is calculated, so as to extract the weighted keywords. According to the keywords, the Chinese texts of the pages visited by the users are classified in order to obtain the user's interest characteristic data. Next, according to the frequency of occurrence of each feature category, the weight of the user's interest feature is calculated. The user's interest field is predicted, and then, resources that match the user's interest field can be pushed to it.

2.3.1. Design of Web Service Architecture. Web service is a distributed computing technology, also known as a platform-independent technology, which is used to integrate and connect applications on various platforms. When the mobile terminal needs to call the service, it needs to find out which server can provide such service on UDDI. UDDI is a technology for describing, discovering, and integrating web service, which is an important part of the web service protocol stack. After finding the server, the mobile terminal asks for the specific calling method, and the server needs to provide the specific method interface of the service according to the method. That is, it returns an XML document described in WSDL format, which describes the interface and its parameters and return type. The mobile terminal uses SOAP to encapsulate the content according to the interface description and sends an HTTP request. Then, the server returns a SOAP packet. In this way, requests and responses at both the server and the mobile end can be unblocked. When the ESB (the full name of ESB is enterprise service bus, which generally refers to enterprise service bus) receives the request message, it first looks up the requested service in the internal service registry and converts the transmission protocol and message format in the request message according to the service needs. Then, it delivers the message to the service provider according to the service path. After obtaining the returned result, the ESB will again

convert the message format or transmission protocol to the type supported by the service requester, finally sending the result to the service requester.

2.3.2. User Login Interface Settings. The user needs to enter the personalized push system through the user interface of the login system, which is the foreground program for the user to perform various operations. Through the research and analysis of the personalized push service objects, the user authentication module of the system is designed according to the level of service required by users. Different levels of users have different needs for services. High-level users may need some highly targeted services, while low-level users may only need some simple and rudimentary services. When designing the landing module, different landing page settings should be made according to the different needs of users. In the system, when the user logs in, a window for whether to customize the service will appear, and the user can choose after authentication. If customized information is required, the system will automatically extract the user's registration information from the user's information database and then match it according to the user's preferences. Then, the user is taken to a page with customized information. If the user does not need to customize the information, the system will automatically enter a relatively simple personalized service interface.

2.3.3. Information Push Interface Design. Personalized push service is a service method with strong initiative and individuality. The information push interface is an important part of the system. This personalized push method is realized by using the relevant theories and technologies of complex networks. The establishment of the user model is dynamic, and the user's information base can be updated in real time with the change in readers' interests. Therefore, the accuracy of the information push is greatly improved. The service quality of the information push is also improved. The specific information push is shown in Figure 2. The information resources to be pushed are clustered according to different subject words. The resources with the same point of interest are divided into one group, and the information resources of different points of interest are divided into different groups. After clustering, several information resource groups with common points of interest within groups and different points of interest among groups are formed.

2.3.4. Data Storage and Interface Design. The services layer is added to the infrastructure to interact directly with the database while reducing the frequency of the controller accessing the model directly, which can reduce coupling and facilitate expansion. In addition, how to choose a good data storage method is very important for an application that needs to display data frequently. Meanwhile, for the interface that frequently displays data, memory management is very important. In addition to using the ARC technology recommended by iOS, the timing of data acquisition and interface refresh, the reuse of cells, and the addition and
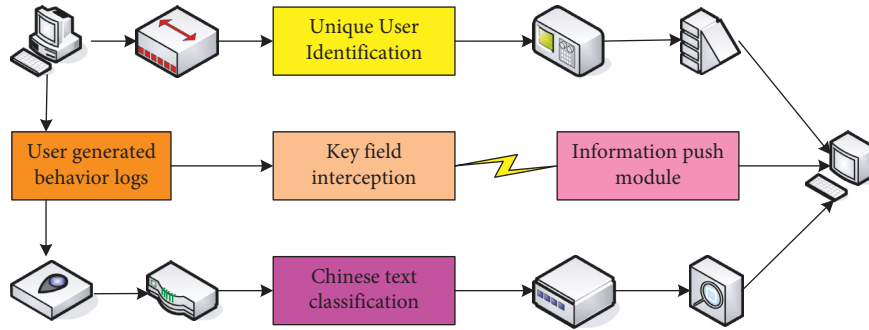
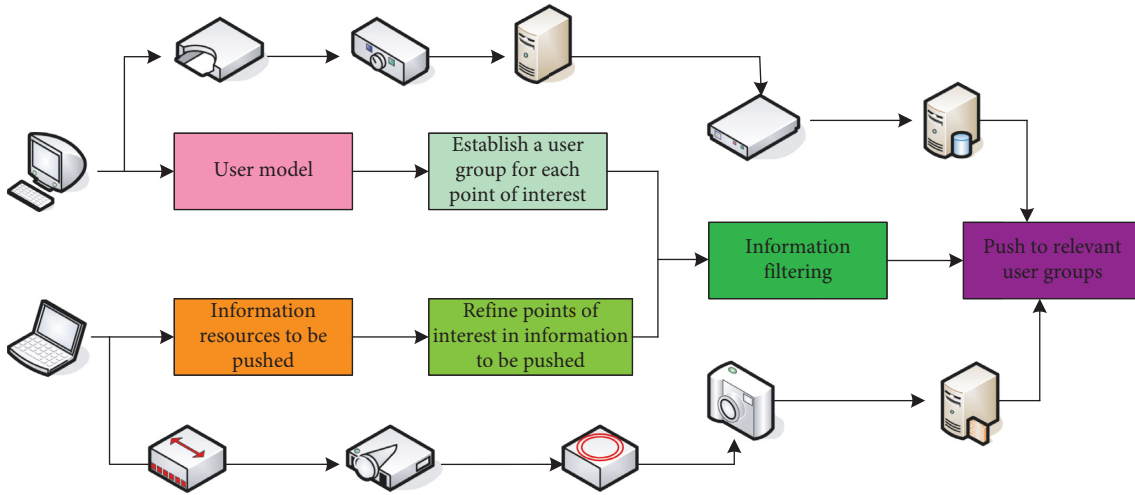FIGURE 1: Architecture design of personalized push service.

FIGURE 2: Specific information push.

deletion of subviews must be carefully grasped. Cell array integrates related data of different types into a single variable, making it easy and convenient to reference and process a large amount of related data.

2.3.5. Design and Implementation of Application Push Module.
The data source of the push module mainly consists of three parts: user personal information, application basic information, and user usage of the application. Collection and processing of user personal information refer that user personal information is obtained through telecom's CRM management system, mainly including the user's name, age, and gender. Large-scale customer relationship management system (CRM) is a computer management system for business processing, operation management, and customer service by telecom key customer management departments. The push module will store this information in a database for later use. The basic information of the application is obtained in the application management module, including the name, category, trigger action, and application icon of the application. In this system, three attributes of application category, delivery time, and trigger action are selected as the characteristics of the project. According to these three attributes, the item attribute matrix is established to calculate the similarity. In terms of measuring user usage

of the application, this paper focuses on how users interact with the application through the application's user interface. According to the RFD model, the user's score for the application is obtained as shown in the following formula:

$$\text{score}_{u,i} = \alpha S_R + \beta S_F + \gamma S_D, \tag{20}$$

where $\text{score}_{u,i}$ represents the user's score for an application (i.e., the user's preference for the application). S_R represents the time difference between the user's last use of the app and the current time. S_F is the number of times the user clicked the app. S_D represents the total time the user interacts with the application.

2.3.6. Design and Implementation of Client Software.
The application package receiver is mainly responsible for monitoring whether the user has installed or uninstalled an application. User behavior monitors are used to record the number and duration of user interactions with the application. The user behavior information uploader will upload the user behavior records to the server at regular intervals. Requests for downloading applications and uploading data are placed in the request queue, and the task executor processes these tasks by issuing HTTP requests to the server. Then, the application converts into the state running in the foreground. If the user starts another application, the current

application loses focus and converts into the background running state. Once the user presses the navigation button or the screen unlock key, the application returns to the foreground running state. When it is terminated by the user or terminated by the android operating system (such as resource recycling), the application will return to the stopped state. If some service center applications may always run in the background, their status will directly change from the stopped state to the background running state.

The personalized push service architecture designed in this paper includes data storage module, information push module, and user login module. The data source of the push module mainly includes three parts: user personal information, application basic information, and user usage of the application. This article focuses on how users interact with an application through the application's user interface.

3. Personalized Push Service Design Results

Data operations such as RSS information storage, user interest model storage, user interest model update storage, and recommendation information storage must be completed by the database. According to the functional requirements of the system, the main database relationships designed using SQL Server 2005 database tools are shown in Table 1.

After receiving the request, the public open platform sends the information to the distribution module. Then, the distribution module will store the information in Table 4-luser__device_token database table. It mainly contains 8 fields, of which the primary key is the unique identifier of the user's device. The user ID is the user's ID information. The application ID is the application ID number applied by the merchant in the management console. The device token is the unique device identifier of the app user's mobile phone. The device type is the operating system type of the user's mobile phone. User extension information is the information customized by the merchant, such as avatar and nickname which is in json format. Creation time is the creation time of this record, and the update time is the update time of this field. The basic information of users is shown in Table 2.

According to user access characteristics, all users can be divided into four types: (1) users with stable interests and narrow interests; (2) users with stable interests and wide interests; (3) users with unstable interests and narrow interests; (4) users whose interests are unstable and wide-ranging. Representative users A, B, C, and D are selected from the four different user types for testing, and the number of users A, B, C, and D using the recommended function is 10. The total amount of information obtained is the same. The test results of the four types of users using the original model are shown in Table 3.

The statistics of interest degree prediction are shown in Figure 3. During the weekdays, the first user was most interested in clothing and digital goods, while on weekends his interest turned to shoes and food. The second user stopped paying too much attention to digital at the weekend, and his interest in home appliances increased greatly. The third user became more interested in food at the weekend and began to

TABLE 1: Main database relationships designed using SQL Server 2005 database tools.

Serial number	Column name	Illustrate
1	Id	ID number of the RSS feed
2	Name	The name of the RSS feed
3	Ur	The address of the RSS feed

TABLE 2: User basic information.

Frequency	Field name	Type
1	id	int
2	user_id	int
3	app_id	varchar (200)
4	device_token	varchar (200)

TABLE 3: Test results of four types of users using the original model.

Frequency	Number of information bodies of interest to users	The number of candidates for the number of information bodies recommended by the system
1	6	8
2	5	9
3	7	9
4	3	3

pay attention to digital goods at the same time. Therefore, the user's interest level is often closely related to the situation. By analyzing the situation, the user's interest can be more accurately grasped, which is also more conducive to the precise marketing of service providers. The error between the interest results under each interest topic in the test set and the results obtained from the previous statistical analysis of the training set is within a reasonable range, and the maximum is about 5%. The accuracy of interest degree prediction in different scenarios can reach more than 90%, which directly confirms the good applicability and effectiveness of the analysis and calculation method and the constructed model of user interest in this paper.

For users with unstable interests, the recall rate is not high; for users with wide interests, the precision rate is not high (the recall rate is shown in Figure 4(a)). The improved model can improve the recall rate and precision rate of the four types of users, and the comprehensive classification rate of the four types of users is also improved, especially for users with unstable interests and wide interests. This shows that the improved method of the user interest model proposed in the study is feasible (the precision rate is shown in Figure 4(b)).

The common push agent stores the crawled news web pages in the local directory corresponding to the source website. Then, the specific web page analysis process analyzes the structure and content of the web page and extracts the news title, text, time, category, and other information. Then, it stores them in the corresponding fields of the news table in the local database. The processing of all web pages relies on a specific analysis of a given web page. The specific
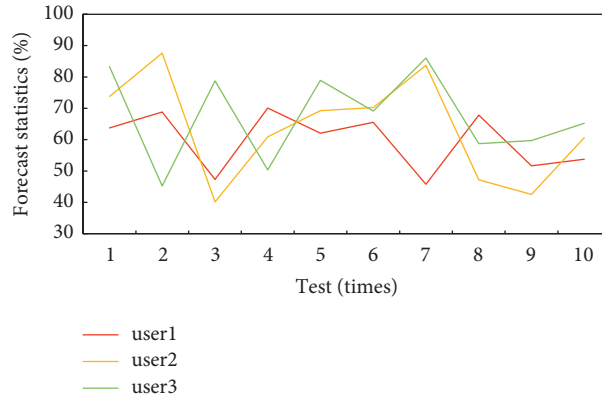
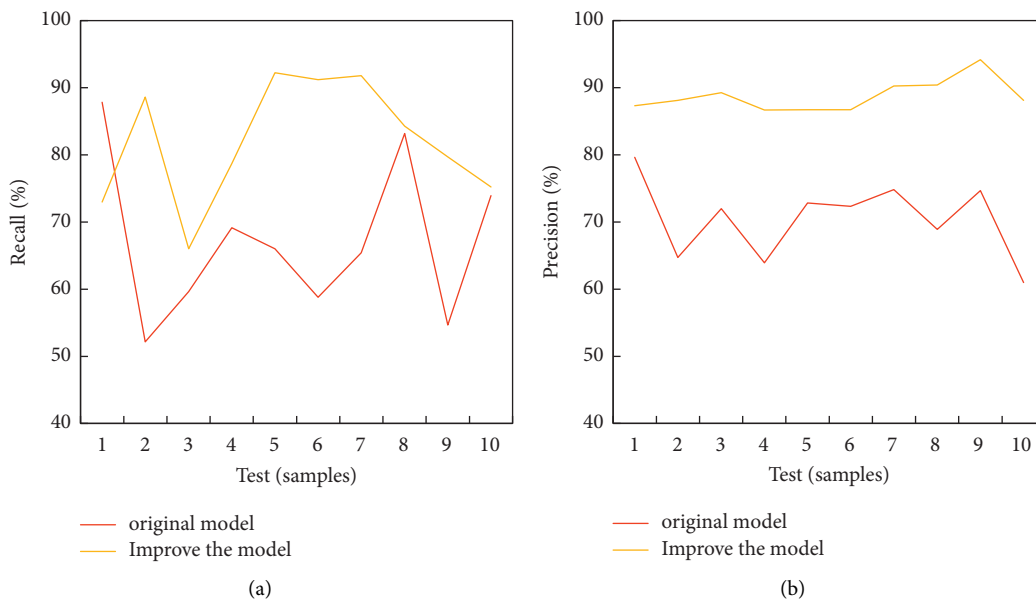FIGURE 3: Interest degree prediction statistics.

FIGURE 4: Comparison of using the original model and the improved model. (a) Recall ratio. (b) Precision ratio.

judgment of the news category is carried out when the news information is stored in the database. The MetaNews relational model is shown in Table 4.

The fixed type contains the least user knowledge and has the lowest weight. The user's direct selection can most clearly reflect the user's tendency and has the highest weight. The weight values of various types of information are shown in Table 5. The reliability coefficient of each field is calculated. When the reliability coefficient of a field exceeds or falls below a certain threshold, the degree of user attention in this field will be changed. For example, when the reliability coefficient of a user's interest field is lower than 0, the field will become the user's uninteresting field.

The smaller the value of MAE is, the better the recommendation effect of the system will be. The user's unrated items are predicted on the test set, so that the number of nearest neighbors gradually increases with a stride of 5 (Pearson correlation and Spearman correlation are shown in Figure 5(a)). Figure 5 shows the MAE values of the four

methods of Pearson correlation, Spearman correlation (it is a nonparametric measure of the dependence of two variables, which uses a monotonic equation to evaluate the correlation of two statistical variables), cosine similarity, and personalized recommendation on different number of neighbors. Among them, the selection of the number of nearest neighbors is 13 cases such as 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, and 65 (cosine similarity and personalized recommendation are shown in Figure 5(b)). In the improved algorithm (personalized recommendation), when the number of neighbors is 40, the value of MAE tends to be stable, and the value of MAE is always smaller than that of the other three similarity measures. The personalized recommendation method proposed in this paper can significantly reduce MAE and improve the recommendation effect of the recommendation system. In subsequent experiments, the number of nearest neighbors k is chosen to be 45.

Except that the result calculated by the recommendation algorithm in the column of accuracy rate is not optimal, the

TABLE 4: MetaNews relational model.

Frequency	Field name	Field type	Notes
1	SiteNO	int	Site ID, also a keyword
2	SiteName	char (100)	Website name
3	IndexURL	char (200)	The address of the website is often also the URL of the news category index page
4	LocalDir	char (200)	The local storage path of the website page

TABLE 5: Various information weight values.

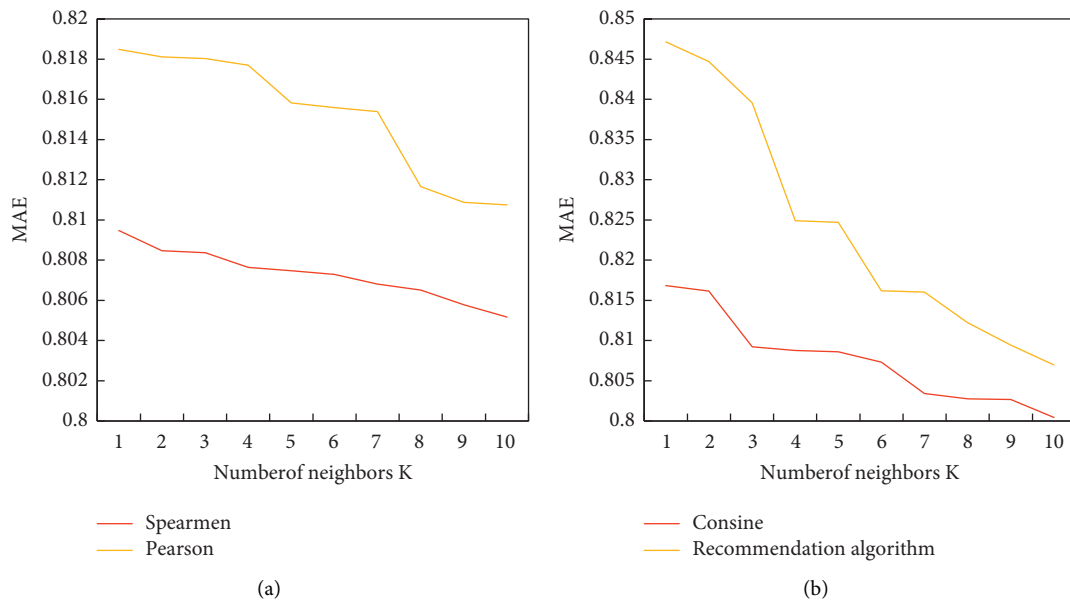| Source	Weights	Certainty
Fixed type	1	1
Simple documentation feedback	3	2
Result documentation feedback	2	2
Direct selection	4	1

FIGURE 5: MAE values under each similarity measure. (a) Pearson correlation and Spearman correlation. (b) Cosine similarity and personalized recommendation.

other two evaluation indicators are optimal. Compared with other similarity measurement methods, it has a relatively obvious improvement (precision and recall rates are shown in Figure 6(a)). The accuracy of the recommendation algorithm is less than one-thousandth lower than the optimal value. Therefore, the recommendation algorithm has a good recommendation effect on the whole (F evaluation is shown in Figure 6(b)).

When the number of neighbors is less than 20, the MAE value of the improved algorithm in this study is smaller than the MAE value obtained by the k-means algorithm and Pearson correlation. When it is greater than 20, the MAE value of the improved algorithm is also greater than that obtained by SVD-based collaborative filtering. In other words, when it is less than 20, the improved algorithm in this study is better than the k-means algorithm and the Pearson-related collaborative filtering recommendation algorithm. When it is greater

than 20, the performance of the improved algorithm in this study is optimal (K-means and Pearson are shown in Figure 7(a)). When the number of neighbors is greater than 28, the improved algorithm in this study has the highest recommendation accuracy among the above four algorithms. When the number of neighbors is lower than 28, in terms of recommendation accuracy, the recommendation accuracy of the improved algorithm is better than that of the K-means algorithm and Pearson-related recommendation, which is slightly lower than the recommendation based on SVD (SVD is singular value decomposition. Its full name is singular value decomposition), but the overall recommendation effect is greater than 80%. It can be seen from the experiments that the bipartite k-means collaborative filtering algorithm based on SVD data dimensionality reduction proposed in this study improves the recommendation effect. Moreover, the good scalability of the algorithm is also
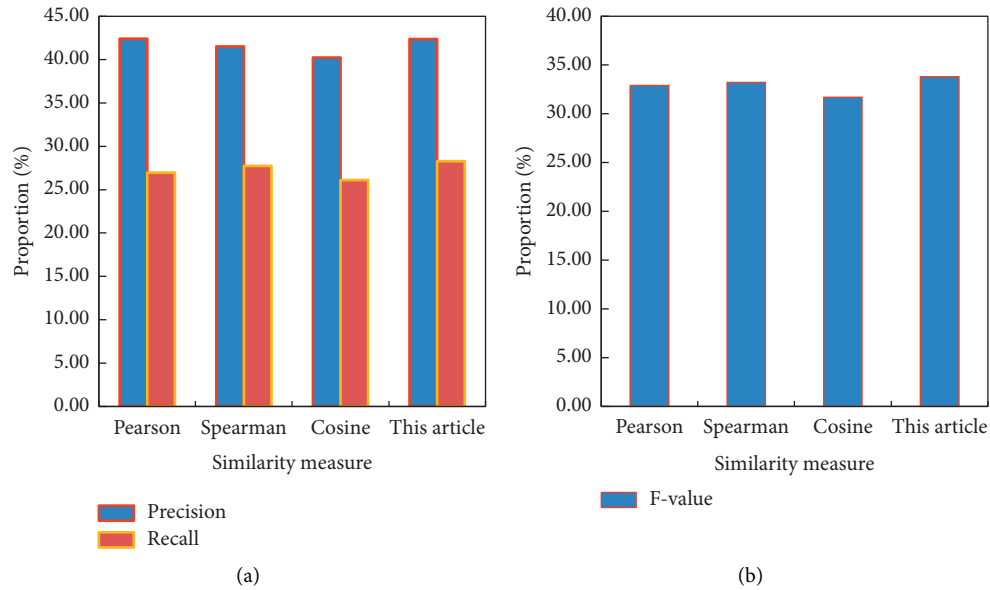
FIGURE 6: Comparison of recommendation quality of different similarity methods. (a) Precision and recall ratio. (b) *F* evaluation.
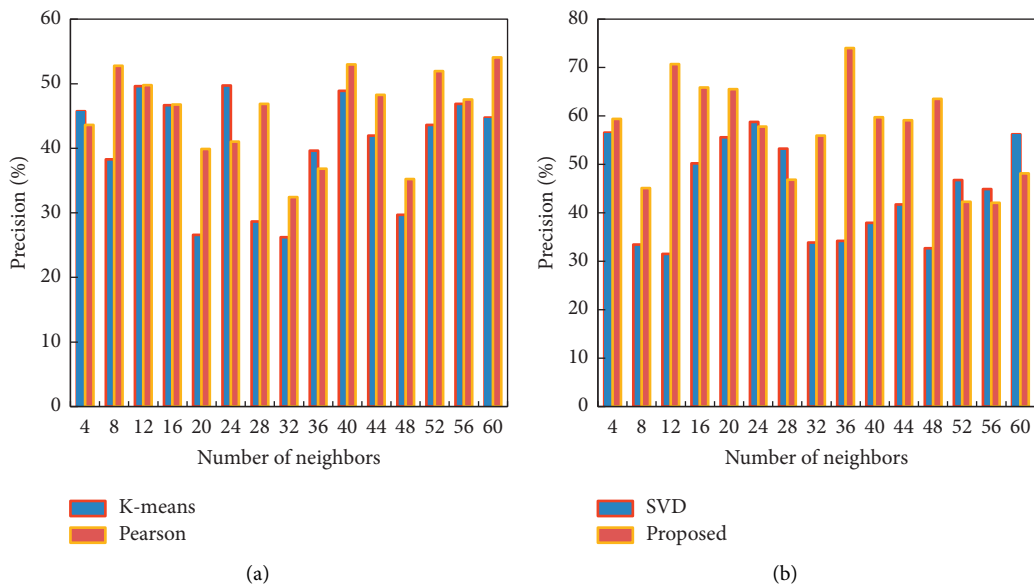

FIGURE 7: Comparison of precision values for different number of neighbors. (a) *K*-means and Pearson. (b) SVD and proposed.

preserved. Meanwhile, the problem of local optima generated by the algorithm is alleviated as much as possible (SVD and proposed are shown in Figure 7(b)).

The test for new users and nonnew users is shown in Figure 8. Some classmates and friends are invited to test the recommendation function of the platform offline. By simulating user registration and logging in to the platform, and entering personal information, accessing part of the website for scoring and other behaviors, the user's personal attributes and historical scoring records are formed in the system. Collection and counting user feedback on the recommended content refers to getting a list of recommended content after logging in to the platform multiple times, as well as testing whether the user is interested in the content in

the list. The test content is aimed at testing the recommendation accuracy of the two recommendation strategies in the combined recommendation algorithm. The test user objects are divided into two cases: new users and nonnew users. A "new user" is a user who does not have any scoring behavior after registering and logging in. "Nonnew users" are users who have generated 20 content rating actions.

In order to compare the performance of hybrid push technology based on interest degree and traditional collaborative filtering technology, the following experimental methods are used in this paper: 10 users are randomly selected as experimental samples, and then, the interest data of 80% of the interest topics of these 10 users are used as the known data of the samples. The interest degree of the
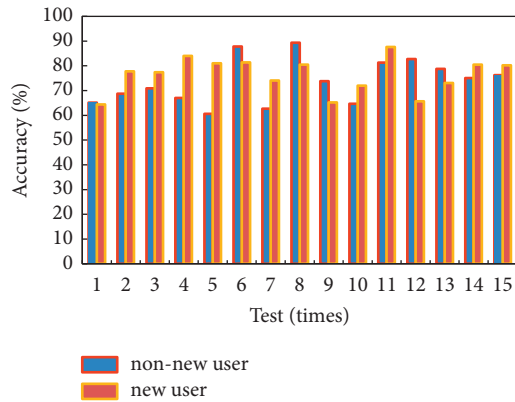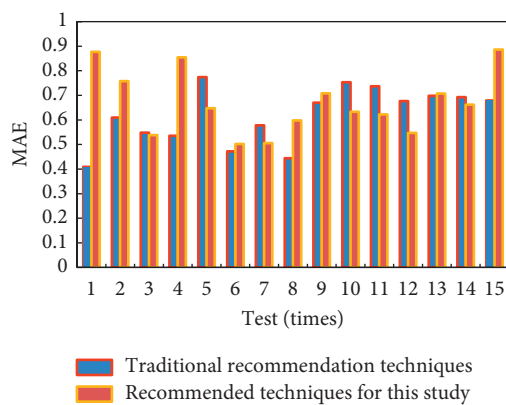
FIGURE 8: New user and nonnew user test.
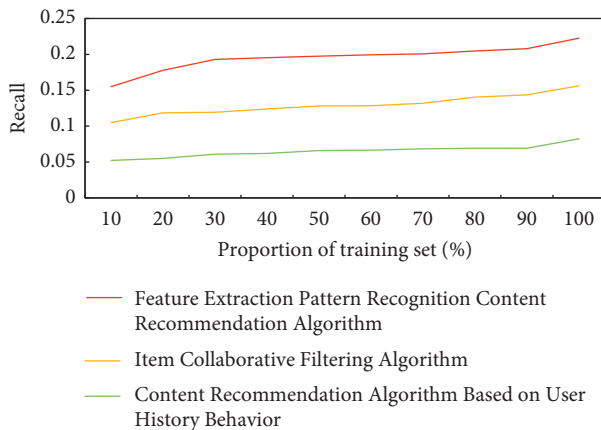

FIGURE 9: MAE value comparison.

FIGURE 10: Comparison of the experimental results of the three algorithms under the hit rate index.

remaining interest topics is predicted by the hybrid push technology based on interest degree and the traditional collaborative filtering technology. Finally, the predicted interest degree is compared with the real interest degree data of the other 20% of interest topics. The comparison of the mean absolute deviation MAE is used to verify the pros and cons of the two push technologies. MAE (the full name of MAE is mean absolute error, that is, the mean absolute value

error, which represents the average value of the absolute error between the predicted value and the observed value) values of the sample users are averaged in each experiment. The MAE value comparison results are shown in Figure 9. The mean absolute deviation MAE directly reflects the level of recommendation accuracy. The experimental results show that the hybrid push technology based on interest degree is better than the traditional collaborative filtering technology in terms of the MAE index, which improves the overall push quality to a certain extent.

The comparison of the experimental results of the three algorithms under the hit rate index is shown in Figure 10. The node data are the mean of the recommendation accuracy of all test users. The hit rate in this paper refers to the ratio between the number of actual "hit" user interests in the content list recommended by the system to the user and the length of the recommended content list. In this experiment, each of the three algorithms recommends 100 pieces of content to each test user (there are more than 9,000 pieces of content in the content dataset). The percentage displayed on the abscissa refers to the proportion of the training set (e.g., 80% means that 80% of the sample data of movie reviews are used for training and 20% is used for testing). In the case of a large number of training samples, the reason for the low hit rate of the three algorithms is that the total number of valid evaluations for a single user is limited (about 100). When the number of training samples is large, the number of test samples is very small, so the number of hit test sets is small. Therefore, when the number of recommended contents remains unchanged (the convention is 100), the hit rate will increase with the reduction of training data. However, in general, the feature extraction pattern recognition content recommendation algorithm based on real-time feedback of user interests proposed in this paper is superior to the traditional mainstream item-based collaborative filtering algorithm (the collaborative filtering algorithm discovers the user's preference based on the mining of the user's historical behavior data. It can predict the products that the user may like to recommend) and user historical behavior-based content recommendation algorithm under different training sets, which is especially obvious when the training samples are few; that is, the data are sparse.

4. Conclusion

Generally speaking, data features are the source of analyzing interest models, which mainly refer to a series of data related to the characteristics and preferences of users. In this paper, a new user interest model is established on the basis of user interest feature information mining, including the user's personal interest set and feature extraction algorithm, so that the personal push agent has the ability to learn and track user interest. The personalized push service system designed in this paper can provide information delivery methods such as e-mail and short messages according to user needs and can be set by users according to their own conditions. The whole system has a common push agent, which is mainly responsible for the operation of various functions at the system level, and interacts, collaborates, and shares knowledge with

the user's personal push agent to jointly complete the task of personalized information push. Due to the limited time, there are still some areas in the research that need to be further studied and improved. The setting of some parameters in the user interest model needs to be further studied, such as the setting of the similarity threshold and the setting of the sliding window size. These parameters have a great influence on the results and should have better adaptability. With the continuous operation of the system, the database will continue to increase. How to better organize and manage information data and optimize the user interest model is the work that needs to be further improved.

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] A. Wohllebe, D. S. Hübner, U. l Radtke, and S. Podruzsik, "Mobile apps in retail: effect of push notification frequency on app user behavior," *Innovative Marketing*, vol. 17, no. 2, pp. 102–111, 2021.

[2] Z. Lyu, Y. Dong, C. l Huo, and W. Ren, "Deep match to rank model for personalized click-through rate prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, pp. 156–163, California USA, Feburary 2020.

[3] C. Xiao-Long, D. Bo, and S. l Guo-Ping, "Multi-collaborative filtering algorithm for accurate push of command information system," *Revista de la Facultad de Ingenieria*, vol. 32, no. 7, pp. 165–172, 2017.

[4] Y. Wang, Y. Tian, X. l Yin, and X. Hei, "A trusted recommendation scheme for privacy protection based on federated learning," *CCF Transactions on Networking*, vol. 3, no. 3-4, pp. 218–228, 2020.

[5] N. Ting, "Yangl. "Incentive mechanisms for data dissemination in autonomous mobile social networks"," *IEEE Transactions on Mobile Computing*, vol. 16, no. 11, pp. 3084–3099, 2017.

[6] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 327–339, 2014.

[7] F. A. S. Borges, R. A. S. Fernandes, I. N. l. Silva, and C. B. S. Silva, "Feature extraction and power quality disturbances classification using smart meters signals," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 824–833, 2016.

[8] R. Matthew, F. Fioranelli, H. Borrion, and M. Griffiths, "Ultistatic micro-Doppler radar feature extraction for classification of unloaded/loaded micro-drones," IET Radar," *Sonar & Navigation*, vol. 11, no. 1, pp. 116–124, 2016.

[9] J. Pei, Y. Huang, W. l Huo, J. Wu, J. Yang, and H. Yang, "SAR imagery feature extraction using 2DPCA-based two-dimensional neighborhood virtual points discriminant embedding," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 6, pp. 2206–2214, 2016.

[10] H. Liu, D. Zhu, S. l Yang et al., "Semisupervised feature extraction with neighborhood constraints for polarimetric SAR classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 7, pp. 3001–3015, 2016.

[11] Y. Zhu and P. Tao, "Design and application of the personalized push system of the agricultural extension service," *Agro Food Industry Hi-Tech*, vol. 28, no. 1, pp. 2690–2694, 2017.

[12] G. Luo, "Personalized online education push model based on user interest in the era of big data," *Revista de la Facultad de Ingenieria*, vol. 32, no. 7, pp. 61–67, 2017.

[13] A. Salavitabar and A. K. Armstrong, "Personalized interventions: a reality in the next 20 Years or pie in the sky," *Pediatric Cardiology*, vol. 41, no. 3, pp. 486–502, 2020.

[14] S. V. Prokopenko, E. Y. Mozheiko, M. V. l Abroskina et al., "Personalized rehabilitation assessment of locomotor functions in Parkinson disease using three-dimensional video analysis of motions," *Russian neurological Journal*, vol. 26, no. 1, pp. 23–33, 2021.

[15] J. X. Liu, D. Wang, Y. L. l Gao et al., "A joint-L2, 1-norm-constraint-based semi-supervised feature extraction for RNA-Seq data analysis," *Neurocomputing*, vol. 228, pp. 263–269, MAR.8.

[16] Y. Shu, "The separation of the interest characteristics of the interpersonal information game in the virtual community," *Psychology of China*, vol. 3, no. 2, pp. 207–217, 2021.

[17] R. Al-Otaibi, N. Jin, T. l Wilcox, and P. Flach, "Feature construction and calibration for clustering daily load curves from smart-meter data," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 645–654, 2016.

[18] C. Aaron, A. Cholaquidis, and A. Cuevas, "Stochastic detection of some topological and geometric feature," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 4596–4628, 2017.

[19] L. Ma, B. Du, H. Chen, and N. Q. Soomro, "Region-of-Interest detection via superpixel-to-pixel saliency analysis for remote sensing image," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1752–1756, 2016.

[20] Y. Xu, W. Xu, and Y. G. l. Ma, "A new study for 16 og; a 12 c at the energies of nuclear astrophysics interest: the inverse of key nucleosynthesis reaction 12 ca; g 16 o," *Nuclear Instruments and Methods in Physics Research A*, vol. 581, no. 3, pp. 866–873, 2017.

Hindawi

Research Article

Feature Recognition and Style Transfer of Painting Image Using Lightweight Deep Learning

Yuanyuan Tan [ID]

College of Art and Design, Hunan First Normal University, Changsha 410205, China

Correspondence should be addressed to Yuanyuan Tan; cylintyy@hnfnu.edu.cn

This work aims to improve the feature recognition efficiency of painting images, optimize the style transfer effect of painting images, and save the cost of computer work. First, the theoretical knowledge of painting image recognition and painting style transfer is discussed. Then, lightweight deep learning techniques and their application principles are introduced. Finally, faster convolutional neural network (Faster-CNN) image feature recognition and style transfer models are designed based on a lightweight deep learning model. The model performance is comprehensively evaluated. The research results show that the designed Faster-CNN model has the highest average recognition efficiency of about 28 ms and the lowest of 17.5 ms in terms of feature recognition of painting images. The accuracy of the Faster-CNN model for image feature recognition is about 97% at the highest and 95% at the lowest. Finally, the designed Faster-CNN model can perform style recognition transfer on a variety of painting images. In terms of style recognition transfer efficiency, the highest recognition transfer rate of the designed Faster-CNN model is about 79%, and the lowest is about 77%. This work not only provides an important technical reference for feature recognition and style transfer of painting images but also contributes to the development of lightweight deep learning techniques.

1. Introduction

With the development of network technology, computer technology has become the mainstay of various industries in the development of human society. As relatively comprehensive computer technology, machine learning, including deep learning technology, is also the focus of current social research. Therefore, deep learning technology has been comprehensively developed in many industries [1]. As an important project in the field of art, painting images have become a research hotspot for image feature recognition and style transfer [2]. The innovation lies in the feature recognition and style transfer research of painting images through lightweight deep learning technology. Although this identification and migration work is not yet mature, many studies have provided technical support.

Shen et al. [3] pointed out that, with the development of computer technology, professional computer technology has gradually emerged and played an important role in the field of image recognition. Therefore, image recognition technology

based on deep learning has been developed to improve image recognition accuracy and has become one of the main technologies that people apply in real life. However, the current network image recognition technology has many shortcomings, including high computational cost, large storage capacity, and unstable algorithms. Maqsood et al. [4] proposed a lightweight image recognition and segmentation technique using depthwise separable convolutions. Its main idea is to realize image distribution recognition through image segmentation. Then, the technique fuses the distributed features of the image. Finally, the comprehensive features of the images are aggregated to improve image recognition efficiency. Ibrahim et al. [5] proposed an improved backpropagation neural network (BPNN) for image feature recognition and learning algorithm and gave the adjustment method of momentum coefficient and learning rate. After research, the accuracy of each index output by the neural network has been significantly improved, and the average recognition accuracy of the microphone image features has reached 92.7%. The recognition speed also meets the requirements of online real-

time detection. Hollandi et al. [6] proposed a concept of image segmentation to realize the conversion of different styles of images. Through image segmentation, the application of image transfer algorithms and style transfer was realized. This method can make the image transfer style according to the need to perform the synchronous transfer of different parts, increase the accuracy of image style transfer, and realize the lightweight optimization of image style transfer. Currently, computers are used to realize image feature recognition and style transfer are relatively advanced methods. However, the implementation of this technology is too much work. Therefore, it is light-provisioned to achieve lightweight technology. Therefore, lightweight deep learning technology is used to realize feature recognition and style transfer of painting images, which is a better method at present.

To sum up, first, the theoretical knowledge of painting image recognition and style transfer is discussed. Then, lightweight deep learning techniques and their application principles are introduced. Finally, through a lightweight deep learning model, faster-convolutional neural network (Faster-CNN) image feature recognition and style transfer models are designed. The innovation is that it can change the traditional method of painting image feature recognition and style transfer and realize the accurate image feature recognition and painting style transfer process under the lightweight deep learning technology. This study not only provides a technical reference for feature recognition and style transfer of painting images but also contributes to the development of lightweight deep learning techniques.

2. Research Theories and Methods

2.1. Feature Recognition of Painting Images. At present, all kinds of science and technology have become the main support point of human life. Therefore, the use of science and technology to transform various fields has become the main driving force for human development [7]. As a common form of information transmission in human life, image acquisition has become one of the mainstream technologies in society. The information image contained in the image can fill some of the shortcomings of human information acquisition technology. Therefore, the importance of image information is self-evident. However, how to accurately obtain the information in the image through science and technology is the main problem at present [8].

As an art form that plays an important role in human development, painting carries the development history of a nation. Additionally, it also reflects the life characteristics of a nation. Therefore, painting is an important part of a national culture [9]. It not only reflects the cultural appearance and aesthetic characteristics of the nation in various forms but also reflects the development process of the nation. It is a unique and important way for human beings to observe and express the world [10]. In the long history of human development, countless paintings and images have been produced. The study of these paintings can not only have a more comprehensive understanding of the development process but also comprehensively understand the development of human civilization and promote the development of human beings [11].

In the long-term process of painting feature recognition, human beings rely the most on digital image recognition technology. This technique can only identify a single feature of an image by its face. As human beings continue to generate comprehensive needs for image features, a single image feature recognition method can no longer meet people's needs [12]. People need more comprehensive image feature recognition technology to explore more comprehensive image information and obtain various information contained in images more comprehensively and accurately [13]. As a relatively mature machine learning technology, deep learning technology can comprehensively extract and recognize image features through deep learning and computational analysis and has become the main technology in image recognition [14]. The specific principles of image recognition technology under different technologies are shown in Figure 1.

In Figure 1, the traditional image recognition technology that relies on digital technology can only recognize a single image feature and cannot recognize other elements in a complex image. Deep learning technology can recognize all the features in the image in all aspects and can comprehensively analyze the complex elements in the image, which has made an important contribution to the development of human image recognition [15].

2.2. Image Style Transfer. Style transfer is the transformation of an ordinary image into another style. Transfer generally includes two types of methods, namely, simulation-based rendering techniques and learning-based rendering techniques. The technology based on simulation drawing focuses on imitating the real painting process and is more suitable for hand-drawn paintings, such as oil paintings or sketches [16]. Learning-based rendering techniques focus on learning the texture features of images and are used in a wider range. The art style is a relatively stable overall artistic feature presented by the interaction between the artist's creative personality and the language and situation of the artwork. In the history of human painting, each drawing may present a different artistic style. As nonart majors, everyone has their own opinions on style. How to change the style of one image to another is even more difficult to define. For programmers, especially machine learning programmers, how to turn an inexplicable thing into an executable program is a problem that plagues many researchers in image style transfer [17]. The specific principle of image style transfer is shown in Figure 2.

Figure 2(a) is the style transfer for fruit painting; Figure 2(b) is the style transfer for train images. Different styles will change the basic form of the image, and the field to which the basic style of the image is applied will also change. Therefore, to meet people's needs for different styles of images, the original images are style-transferred to make certain changes according to different style standards to meet different application needs. In the long-term development of deep learning technology, through image recognition and image feature extraction, according to the image style template, the technology of image-oriented style
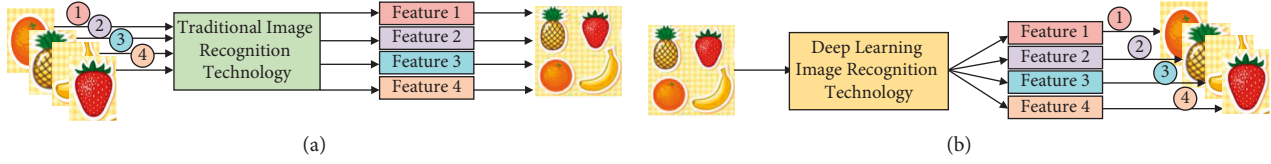
FIGURE 1: The specific principles of image recognition technology under different technologies. (a) Traditional image recognition. (b) Deep learning technology image recognition.
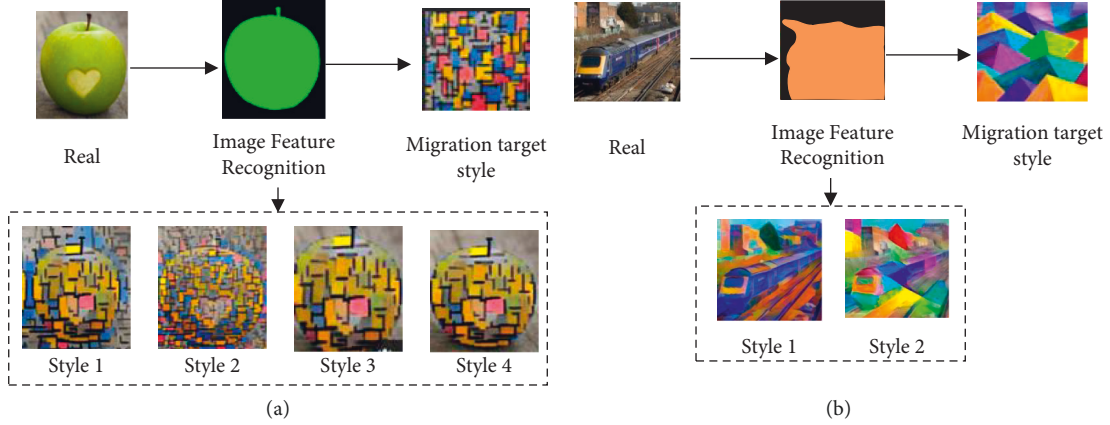

FIGURE 2: The transfer principle of image style. (a) Fruit. (b) Training.

transfer has become the current mainstream technology after long-term development. Therefore, the study of image style transfer through deep learning technology can not only meet the needs of human beings for different styles of images but also comprehensively promote the development and application of deep learning technology [18].

2.3. Lightweight Deep Learning Technology. Since the end of the twentieth century, artificial neural network (ANN) technology has been greatly developed due to its unique properties. The design of this technology mainly refers to the human nervous system. That is, a neural network layer with different properties is constructed through many neurons. Then, a huge neural network system is built from these network layers. Finally, the constructed neural network system can simulate the analysis properties of the human neural network system to a great extent. Therefore, the style exhibited by ANN when dealing with tasks is very human-friendly [19]. The basic structure of neural network technology includes input, hidden, and output layers. Among them, the hidden layer is the main computing level of neural network technology, and its network structure is very deep.

Among neural network technologies, convolutional neural network (CNN) technology is a neural network that focuses on processing images. The technology has evolved considerably since its original design. Therefore, CNN has made an important contribution to humans in the process of processing image information. However, the computationally heavy process of this technology also limits its application in various fields. Therefore, reducing the computational cost of CNN and shortening the computation time, comprehensively optimizing the technology, and

highlighting its contribution to image recognition technology are the main research purposes in the current image recognition field [20].

CNN is a kind of neural network that imitates the visual structure of biology, and it is an efficient recognition algorithm. This recognition algorithm mainly includes convolutional, pooling, and fully connected layers. The convolution operation of the convolutional layer is one of the main operations of CNN. The convolution calculation of continuous functions is shown as follows:

$$\mathbf{s}(\mathbf{t}) = \int \mathbf{x}(\mathbf{a})\mathbf{w}(\mathbf{t}-\mathbf{a})\mathbf{da}. \tag{1}$$

In (1), $\mathbf{x}$ and $\mathbf{w}$ represent integrable functions; $\mathbf{a}$ and $\mathbf{t}$ represent different computational elements; $\mathbf{d}$ represents the convolution operation. The convolution calculation of discrete functions is shown as follows:

$$\mathbf{s}(\mathbf{n}) = \sum_{\mathbf{m}} \mathbf{r}(\mathbf{m})\mathbf{v}(\mathbf{n}-\mathbf{m}). \tag{2}$$

In (2), $\mathbf{r}$ and $\mathbf{v}$ denote discrete functions; $\mathbf{m}$ and $\mathbf{n}$ denote calculation elements [21]. In computer vision tasks, convolution can be regarded as a filtering operation. Usually, a two-dimensional image is used as input data. A two-dimensional discrete convolution is used during convolution as follows:

$$\mathbf{I}(\mathbf{x},\mathbf{y}) * \mathbf{k}(\mathbf{x},\mathbf{y}) = \sum_{\mathbf{s}=0}^{\mathbf{m}} \sum_{\mathbf{t}=0}^{\mathbf{n}} \mathbf{k}(\mathbf{s},\mathbf{t})\mathbf{I}(\mathbf{x}-\mathbf{s},\mathbf{y}-\mathbf{t}). \tag{3}$$

In (3), $\mathbf{I}$ represents the output feature; $\mathbf{k}$ represents the convolution kernel; $\mathbf{m}$ and $\mathbf{n}$ represent the dimensions of the

convolution kernel; $\mathbf{x}$ and $\mathbf{y}$ represent the point of feature output; $\mathbf{s}$ and $\mathbf{t}$ represent the feature extraction point. The functions of the pooling and the fully connected layer are to pool the image and output the result, respectively. The computation of the CNN model includes forward and backward propagation. Forward propagation is a series of computing operations such as image recognition and image feature extraction through input data, and the results are integrated and output. Backpropagation refers to the input of the calculation results to calculate the error as the basic reference information for model optimization. Through continuous iterative training and updating, the parameters learned by the network are optimized, and the training is terminated when the artificially set conditions are reached [22]. Among them, the calculation process of backpropagation is to forward the input sample $(\mathbf{x}, \mathbf{y})$ to calculate the output value of $\mathbf{L_1}, \mathbf{L_2}, \ldots, \mathbf{L_n}$ and the error of the output layer as follows:

$$\boldsymbol{\delta}_{\mathbf{i}}^{(\mathbf{n_i})} = -\left(\mathbf{y} - \mathbf{a}^{(\mathbf{n_i})}\right) \cdot \mathbf{f}'\left(\mathbf{z}_{\mathbf{i}}^{(\mathbf{n_i})}\right). \tag{4}$$

The error calculation of each layer is shown as follows:

$$\boldsymbol{\delta}^{(l)} = \left(\left(\mathbf{W}^{(l)}\right)^{\mathbf{T}} \boldsymbol{\delta}^{(l+1)}\right) \mathbf{f}'\left(\mathbf{z}^{(l)}\right). \tag{5}$$

The partial derivatives of weights and biases are calculated as follows:

$$\nabla_{\mathbf{w}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{x}, \mathbf{y}) = \boldsymbol{\delta}^{(l+1)} \left(\mathbf{a}^{(l)}\right)^{\mathbf{T}}, \tag{6}$$

$$\nabla_{\mathbf{b}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{x}, \mathbf{y}) = \boldsymbol{\delta}^{(l+1)}. \tag{7}$$

The updated weight parameters are as follows:

$$\mathbf{W}' = \mathbf{W}^l - \boldsymbol{\mu} \nabla_{\mathbf{W}^{(j)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{x}, \mathbf{y}), \tag{8}$$

$$\mathbf{b}' = \mathbf{b}^l - \boldsymbol{\mu} \nabla_{\mathbf{b}^{(i)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{x}, \mathbf{y}). \tag{9}$$

δ represents the difference between the true and the predicted value of the network; $\mathbf{W}$ represents the weight; $\mathbf{b}$ represents the bias of the neuron; $\mathbf{z}$ represents the input of the neuron; $\mathbf{a}$ represents the output of the neuron; $\mathbf{f}'\left(\mathbf{z}^{(l)}\right)$ represents the activation function; μ represents the learning rate; l represents the level of neurons; $\mathbf{i}$ represents neurons; $\mathbf{T}$ represents a constant. The loss function of a sample is calculated as follows:

$$\mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{x}, \mathbf{y}) = \frac{1}{2} \left\| \mathbf{y} - \mathbf{h}_{\mathbf{W}', \mathbf{b}}(\mathbf{x}) \right\|^2. \tag{10}$$

The staggered computation of neurons is shown as follows:

$$\begin{aligned} \boldsymbol{\delta}_{\mathbf{i}}^{(\mathbf{n_l})} &= \frac{\partial}{\partial \mathbf{z}_{\mathbf{i}}^{(\mathbf{n_l})}} \frac{1}{2} \left\| \mathbf{y} - \mathbf{h}_{\mathbf{w}, \mathbf{b}}(\mathbf{x}) \right\|^2 \\ &= -\left(\mathbf{y}_{\mathbf{i}} - \mathbf{a}_{\mathbf{i}}^{(\mathbf{n_i})}\right) \cdot \mathbf{f}'\left(\mathbf{z}_{\mathbf{i}}^{(\mathbf{n_f})}\right), \end{aligned} \tag{11}$$

$$\mathbf{h}_{\mathbf{w}, \mathbf{b}}(\mathbf{x}) = \mathbf{a}_{\mathbf{i}}^{(\mathbf{n_i})} = \mathbf{f}\left(\mathbf{z}_{\mathbf{i}}^{(\mathbf{n_i})}\right). \tag{12}$$

To reduce storage space and computational consumption, the CNN model v is compressed. Here, the compression processing method is a matrix decomposition method. This method incorporates the singular value decomposition (SVD) algorithm. The SVD algorithm is a very important model compression method. Its connotation is to represent the original matrix by extracting key features; that is, a complex matrix is approximated by multiplying several small matrices representing key features in the original matrix [23]. SVD is an algorithm widely used in the field of machine learning. It can be used not only for feature decomposition in dimensionality reduction algorithms, but also for recommendation systems and natural language processing. It is the cornerstone of many machine learning algorithms. As a very basic algorithm, SVD has its presence in many machine learning algorithms. In the current era of big data, SVD has a wide range of applications due to its parallelization. The disadvantage of SVD is that the decomposed matrix is not interpretable, but this disadvantage does not affect its use. Singular value decomposition is an algorithm that can be applied to any matrix decomposition; for example, let the data input to the fully connected layer be of size $\mathbf{u} \times \mathbf{v}$ and the weight matrix be W. The calculation of the output data of the fully connected layer is shown as follows:

$$\mathbf{y} = \mathbf{W_x}. \tag{13}$$

$\mathbf{W}$ is used to perform SVD, and the decomposition of $\mathbf{W}$ is replaced by the first $\mathbf{t}$ important eigenvalues after decomposition as follows:

$$\mathbf{W} = \mathbf{U} \sum \mathbf{V}^{\mathbf{T}} \approx \mathbf{U} \sum_{\mathbf{t}} \mathbf{V}^{\mathbf{T}}. \tag{14}$$

$\mathbf{U}$ represents a $\mathbf{u} \times \mathbf{t}$-dimensional orthogonal matrix; $\sum$ represents a diagonal matrix; $\mathbf{V}$ represents a $\mathbf{v} \times \mathbf{t}$-dimensional orthogonal matrix. Therefore, the representation of SVD is as follows:

$$\mathbf{y} = \mathbf{Wx} \approx \mathbf{U} \cdot \left(\sum_{\mathbf{t}} \mathbf{V}^{\mathbf{T}}\right) \cdot \mathbf{x} = \mathbf{U} \cdot \mathbf{z}. \tag{15}$$

The SVD algorithm can decompose the CNN technology, which greatly reduces the computational load of the network. This method is not only simple but also achieves better results. Another method is the low-rank decomposition (LRD) algorithm. This algorithm is an optimization algorithm for the SVD. The SVD algorithm cannot solve the simpler computational operations of images, and the LRD algorithm can solve this problem well [24]. The definition of its output feature map is shown as follows:

$$\mathbf{F_n}(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{i}=1}^{\mathbf{C}} \sum_{\mathbf{x}'=1}^{\mathbf{X}} \sum_{\mathbf{y}'=1}^{\mathbf{Y}} \mathbf{Z}^{\mathbf{c}}(\mathbf{x}', \mathbf{y}') \mathbf{W}_{\mathbf{n}}^{\mathbf{c}}(\mathbf{x} - \mathbf{x}', \mathbf{y} - \mathbf{y}'). \tag{16}$$

$\mathbf{W}$ represents the channel, $\mathbf{n}$ represents the filter, and $\mathbf{C}$ represents the channel's position. The main purpose is to find an approximation of $\mathbf{W}$, $\widehat{\mathbf{W}}$ as follows:

$$\widehat{W}_{\mathbf{n}}^{\mathbf{c}} = \sum_{\mathbf{k}=1}^{\mathbf{K}} \mathbf{H}_{\mathbf{n}}^{\mathbf{k}} \left(\mathbf{V}_{\mathbf{k}}^{\mathbf{c}}\right)^{\mathbf{T}}. \tag{17}$$

K represents the hyperparameter that controls the rank; **H** represents the horizontal filter; **V** is the vertical filter. However, LRD also has certain disadvantages. That is, although LRD has achieved good results for model compression and acceleration, the implementation of this method is not easy. There are decomposition operations with high computational costs. Since different layers contain different information, it is not possible to use a global variable to implement LRD, and it is necessary to perform low-rank approximation (LRA) layer by layer. Moreover, after being decomposed, a lot of fine-tuning training is required to make the network converge to achieve the optimal effect [25].

2.4. Optimized Lightweight Faster-CNN Algorithm.

Machine learning has developed for a long time since its inception, and it has accumulated many shortcomings. Traditional machine learning techniques require continuous design by humans to gradually improve their own learning process. Therefore, in the calculation process, its dependence is particularly large, and the basic technical ability of the operator is relatively high. The development of machine learning has also been greatly limited. Meanwhile, the algorithm's accuracy for image recognition is very low, and it cannot quickly achieve accurate image recognition [26]. Among them, the most important is that the traditional machine learning technology cannot accurately identify various factors in the image, which usually causes large errors in the application. The most serious is that machine learning cannot classify the main part and background part of the image and cannot identify the main information contained in the image [27]. The optimized deep learning technology solves the shortcomings of traditional machine learning technology in image recognition. If deep learning image recognition technology needs to be pushed to a wider field, it is necessary to optimize its calculation process to reduce its computational cost and improve its computational efficiency. Therefore, by optimizing the CNN technology and generating the Faster-CNN model, the model calculation process is simpler, and the calculation effect is also improved to a certain extent. The basic idea of the designed lightweight Faster-CNN model is shown in Figure 3.

In Figure 3, when the Faster-CNN model is used for image feature recognition, it can not only greatly reduce the process of image feature recognition and improve the efficiency of image feature recognition but also optimize the recognition effect of the model and better help people to complete the style transfer of painting images. The approach adopted when optimizing the Faster-CNN model is the region proposal network. It adjusts and optimizes the image recognition area of the Faster-CNN model through anchor points as follows:

$$\mathbf{x} = \mathbf{w_a t_x} + \mathbf{x_a}, \tag{18}$$

$$\mathbf{y} = \mathbf{h_a t_y} + \mathbf{y_a}, \tag{19}$$

$$\mathbf{w} = \mathbf{w_a} \left(\mathbf{t_w} \right), \tag{20}$$
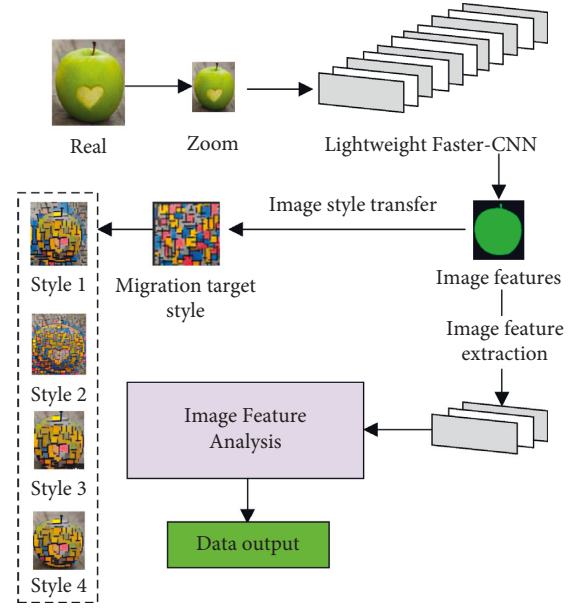
FIGURE 3: The basic idea of the lightweight Faster-CNN model.

$$\mathbf{h} = \mathbf{h_a} \left(\mathbf{t_h} \right). \tag{21}$$

$\mathbf{x_a}$, $\mathbf{y_a}$, $\mathbf{w_a}$, and $\mathbf{h_a}$, respectively, represent the abscissa and ordinate of the center point of the anchor point and the width and height of the anchor point. $\mathbf{x}$, $\mathbf{y}$, $\mathbf{w}$, and $\mathbf{h}$ represent the horizontal and vertical coordinates of the center selected by the model, as well as the selected width and height, respectively. $\mathbf{t}$ represents the corrected value.

2.5. Description of Study Data.

Public datasets are used to train the model comprehensively. The comprehensive performance of the model is evaluated. Here, the adopted dataset includes the Mixed National Institute of Standards and Technology (MNIST) dataset. The MNIST dataset is one of the most popular deep learning datasets. It is a dataset of handwritten digits containing a training set of 60,000 examples and a test set of 10,000 examples. This is a great database for trying out learning techniques and deeply identifying patterns in real data that can spend minimal time and effort in data preprocessing. The open images dataset is a dataset of nearly nine million images of uniform resource locations (URLs). These images span thousands of class image-level label bounds and are annotated. The dataset contains a training set of 9,011,219 images, a validation set of 41,260 images, and a test set of 125,436 images. The Canadian Institute For Advanced Research-10 (CIFAR-10) dataset is another dataset for image classification. It consists of 60,000 images of 10 classes (each class is represented as a row in the image above). There are 50,000 training and 10,000 test images in total. The dataset is divided into six parts: five training batches and one testing batch. Each batch has 10,000 images. The ImageNet dataset is an image dataset organized based on the WordNet hierarchy. WordNet contains about 100,000 phrases, and ImageNet provides an average of about 1000 images to illustrate each phrase. The
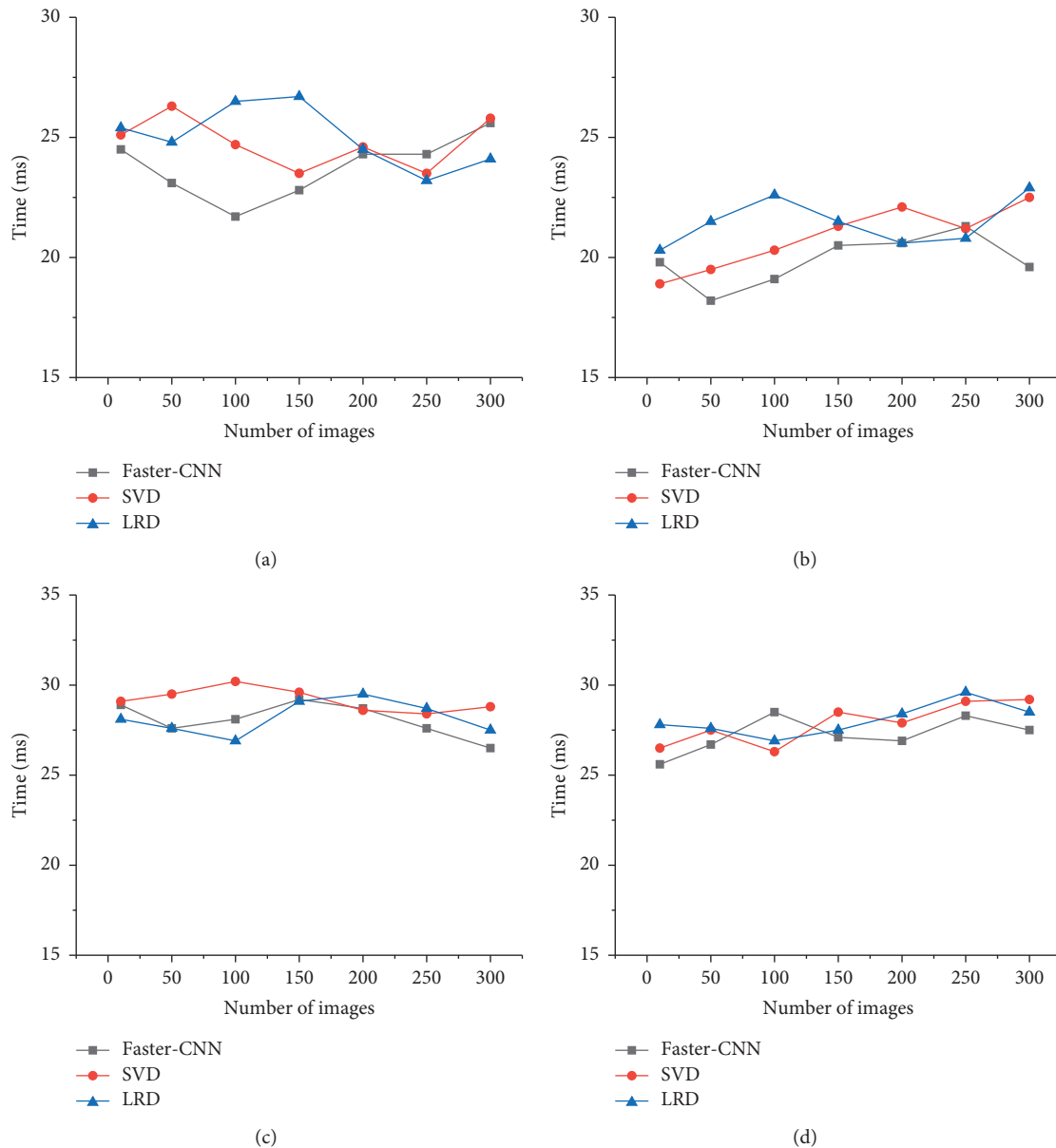
(a)

(b)

(c)

(d)

FIGURE 4: The lightweight deep learning model for image recognition rate. (a) MNIST dataset. (b) Open images dataset. (c) CIFAR-10 dataset. (d) ImageNet dataset.

total number of images is about 1,500,000. Each image has multiple bounding boxes and corresponding class labels.

3. Drawing Image Feature Recognition and Evaluation of Style Transfer

3.1. Identification and Evaluation of Image Features. To improve the computing efficiency and save computing costs and storage space of the deep learning technology, the lightweight deep learning model can be optimized to meet the demand for image feature recognition using the lightweight deep learning model. The average rate of image recognition of the designed model is shown in Figure 4.

In Figure 4, in the image recognition rate evaluation of the lightweight deep learning model, the designed Faster-

CNN model has little difference between the image recognition rates of the SVD and LRD algorithms. The designed Faster-CNN model has the highest average recognition time of about 28 ms in the four datasets, and the lowest is about 17.5 ms. The average image recognition time of the SVD algorithm is around 31 ms at the highest and around 18 ms at the lowest. The average image recognition time of the LRD algorithm is around 30 ms at the highest and around 20 ms at the lowest. The designed model still has advantages over SVD and LRD algorithms. While focusing on the recognition efficiency of the model, it is generally necessary to evaluate the accuracy of the model to determine its comprehensive advantages. The results of the image recognition accuracy evaluation of the model are shown in Figure 5.
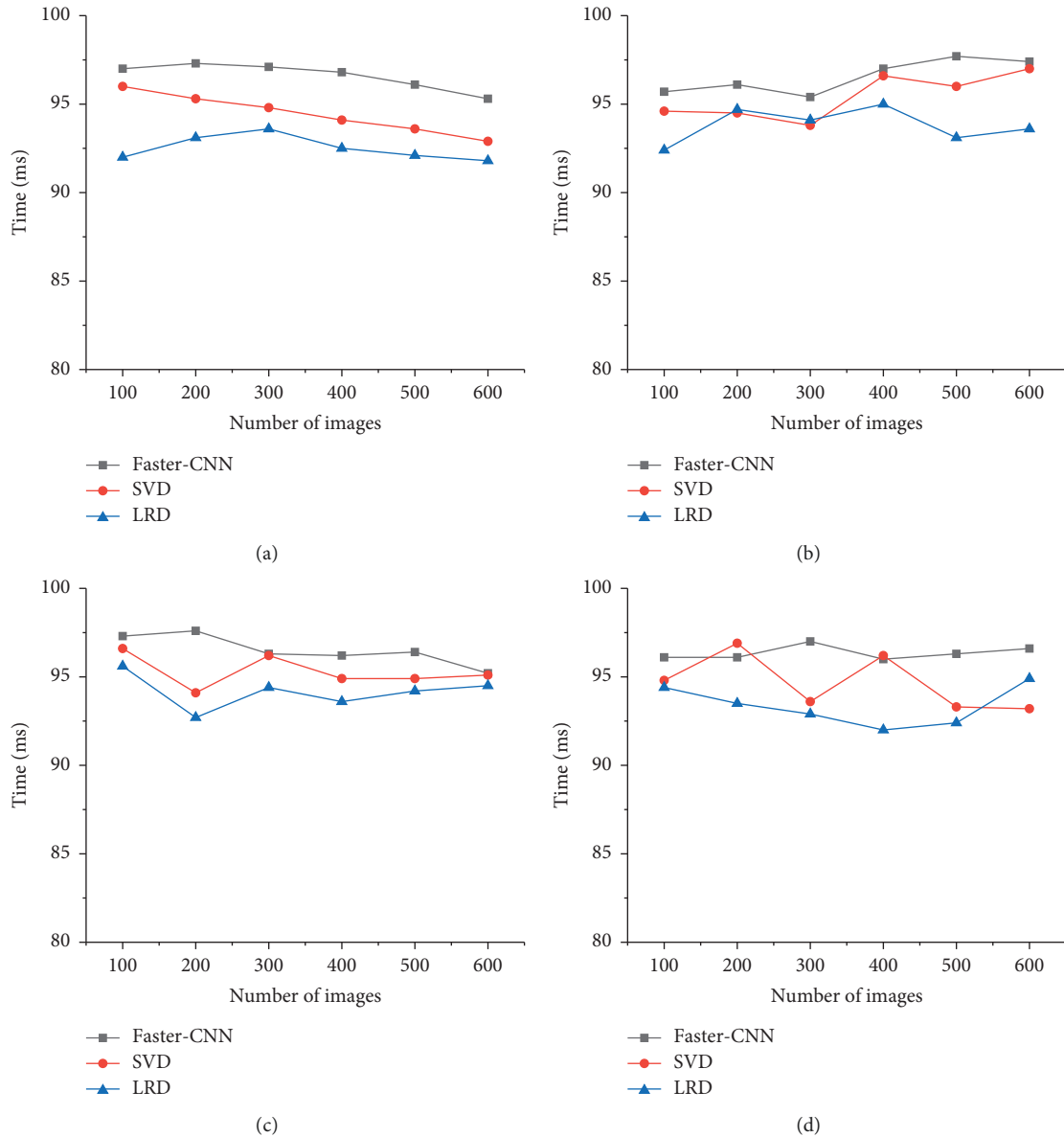
FIGURE 5: The image recognition accuracy of lightweight deep learning models. (a) MNIST dataset. (b) Open images dataset. (c) CIFAR-10 dataset. (d) ImageNet dataset.

In Figure 5, the image recognition accuracy of the designed Faster-CNN model is around 97% at the highest and around 95% at the lowest. The image recognition rate of the SVD algorithm is around 97% at the highest and around 93% at the lowest. The image recognition rate of LRD is around 95% at the highest and around 92% at the lowest. The Faster-CNN model has great advantages in image recognition accuracy.

3.2. Transfer Evaluation of Painting Image Styles. With the development of science and technology, the style transfer of painting images has become a technology that can be generally realized. However, its efficiency and accuracy also need to be improved to a certain extent to achieve an efficient

transfer of painting image styles. The results of images being transferred with different styles are shown in Figure 6.

In Figure 6, the Faster-CNN model can identify and transfer styles of various painting images, so that the styles of the images are transformed into images of other styles. The designed model also has great advantages in the transfer efficiency of image style recognition. The evaluation results of the transfer efficiency of painting image recognition of the Faster-CNN model are shown in Figure 7.

In Figure 7, the Faster-CNN model has a great advantage in transferring painting image style recognition. Among the five groups of image style recognition migration, the recognition migration rate of the Faster-CNN model is the highest at around 79% and the lowest at around 77%. The SVD algorithm recognizes that the highest mobility rate is

FIGURE 6: The transfer results of the recognition style of the painting image of the lightweight deep learning model. (a) Stone tools. (b) Architecture and landscape. (c) Landscape. (d) Pure architecture. Image source is the network.
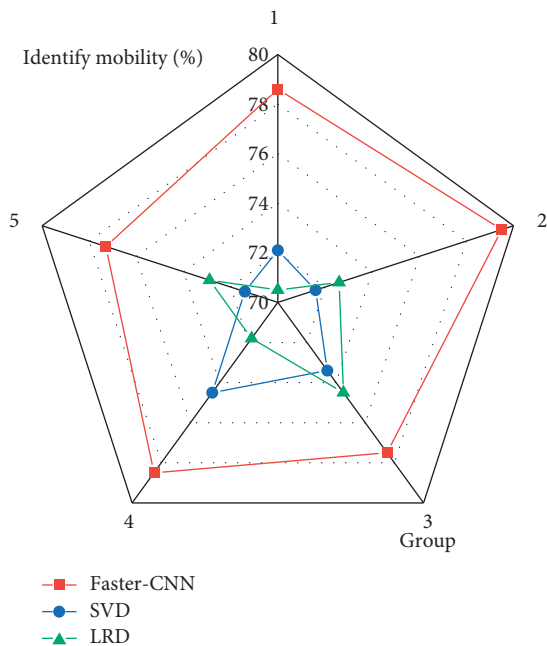

FIGURE 7: Evaluation of transfer efficiency for style recognition in painting images.

around 74%, and the lowest is around 71%. The LRD algorithm recognizes that the highest mobility is around 74%, and the lowest is around 70%. The designed Faster-CNN model still has a strong advantage in the style transfer of painting images.

4. Conclusion

With the development of science and technology, deep learning technology has become a relatively advanced computer technology. The lightweight deep learning model has been greatly optimized in terms of storage and computation based on inheriting the deep learning model technology. First, the theoretical knowledge of painting image recognition and style transfer is discussed. Then, lightweight deep learning techniques and their application principles are introduced. Finally, based on the lightweight deep learning model, Faster-CNN image feature recognition and style transfer models are designed. Model performance is comprehensively evaluated. Studies showed that the designed Faster-CNN model has the highest average recognition time of about 28 ms and the lowest of about 17.5 ms in feature recognition of painting images. The accuracy of

the Faster-CNN model for image feature recognition is about 97% at the highest and 95% at the lowest. Finally, the designed Faster-CNN model can perform style recognition transfer on various painting images. In terms of style recognition transfer efficiency, the highest recognition transfer rate of the designed Faster-CNN model is about 79%, and the lowest is about 77%. The innovation lies in transforming traditional image feature recognition and style transfer methods and realizing painting image feature recognition and style transfer technology under the lightweight deep learning technology. This study designs a model with painting image recognition and style transfer, but the shortcomings of the model have not been comprehensively studied. Therefore, future research will strengthen the development of model defects to optimize and improve the model performance continuously.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] E. Goceri, "Diagnosis of skin diseases in the era of deep learning and mobile technology," *Computers in Biology and Medicine*, vol. 134, no. 9, Article ID 104458, 2021.

[2] X. Wang, L. T. Yang, L. Song, H. Wang, L. Ren, and M. J. Deen, "A tensor-based multiattributes visual feature recognition method for industrial intelligence," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2231–2241, 2021.

[3] J. Shen, N. Liu, and H. Sun, "Vehicle detection in aerial images based on lightweight deep convolutional network," *IET Image Processing*, vol. 15, no. 1, pp. 479–491, 2021.

[4] M. Maqsood, I. Mehmood, R. Kharel, K. Muhammad, J. Lee, and W. Alnumay, "Exploring the role of deep learning in industrial applications: a case study on coastal crane casting recognition," *Hum. Cent. Comput. Inf. Sci*, vol. 11, no. 2, pp. 1–14, 2021.

[5] A. D. Ibrahim, H. M. A. Hussein, and S. A. Abdelwahab, "Automatic feature recognition of cross holes in hollow cylinders," *Journal of the Institution of Engineers: Series C*, vol. 102, no. 2, pp. 257–274, 2021.

[6] R. Hollandi, A. Szkalisity, T. Toth et al., "nucleAIzer: a parameter-free deep learning framework for nucleus segmentation using image style transfer," *Cell Systems*, vol. 10, no. 5, pp. 453–458.e6, 2020.

[7] . Yunfei, . Duzhou, . Li, and . Baocheng, "Lightweight feature fusion network design for local feature recognition of noncooperative target," *Infrared and Laser Engineering*, vol. 49, no. 7, Article ID 20200170, 2020.

[8] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face recognition systems: a survey," *Sensors*, vol. 20, no. 2, p. 342, 2020.

[9] I. J. Jacob and P. E. Darney, "Design of deep learning algorithm for IoT application by image based recognition," *Journal of ISMAC*, vol. 3, no. 3, pp. 276–290, 2021.

[10] S. Knez and L. Šajn, "Food object recognition using a mobile device: evaluation of currently implemented systems," *Trends in Food Science & Technology*, vol. 99, no. 3, pp. 460–471, 2020.

[11] G. Lou and H. Shi, "Face image recognition based on convolutional neural network," *China communications*, vol. 17, no. 2, pp. 117–124, 2020.

[12] H. Ma, N. Hu, and C. Fang, "The biometric recognition system based on near-infrared finger vein image," *Infrared Physics & Technology*, vol. 116, Article ID 103734, 2021.

[13] F. Özyurt, "Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures," *The Journal of Supercomputing*, vol. 76, no. 11, pp. 8413–8431, 2020.

[14] H. Li, C. Huang, and L. Gu, "Image pattern recognition in identification of financial bills risk management," *Neural Computing & Applications*, vol. 33, no. 3, pp. 867–876, 2021.

[15] T. Tsuji, K. Sato, and S. Sakaino, "Contact feature recognition based on MFCC of force signals," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5153–5158, 2021.

[16] C. T. Lin, S. W. Huang, Y. Y. Wu, and S. H. Lai, "GAN-based day-to-night image style transfer for nighttime vehicle detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 951–963, 2021.

[17] C. Zhu, W. Yan, X. Cai, S. Liu, T. H. Li, and G. Li, "Neural saliency algorithm guide bi-directional visual perception style transfer," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 1–8, 2020.

[18] C. Peng, N. Wang, J. Li, and X. Gao, "Universal face photo-sketch style transfer via multiview domain translation," *IEEE Transactions on Image Processing*, vol. 29, no. 8, pp. 8519–8534, 2020.

[19] H. C. Dan, G. W. Bai, and Z. H. Zhu, "Application of deep learning-based image recognition technology to asphalt–aggregate mixtures: Methodology," *Construction and Building Materials*, vol. 297, no. 6, Article ID 123770, 2021.

[20] Z. Yu, W. Chen, and T. Zhang, "Motor imagery EEG classification algorithm based on improved lightweight feature fusion network," *Biomedical Signal Processing and Control*, vol. 75, no. 3, Article ID 103618, 2022.

[21] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, p. 420, 2021.

[22] Y. Guo, B. Pan, Y. Fu, and M. Meng, "CAM-FoC: a high accuracy lightweight deep neural network for grip force measurement of elongated surgical instrument," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. 8, pp. 1–12, 2021.

[23] A. Wani, R. Khaliq, and R. Khaliq, "SDN-based intrusion detection system for IoT using deep learning classifier

(IDSIoT-SDL),” *CAAI Transactions on Intelligence Technology*, vol. 6, no. 3, pp. 281–290, 2021.

[24] L. Deng, J. Li, and Z. Han, “Online defect detection and automatic grading of carrots using computer vision combined with deep learning methods,” *Lebensmittel-Wissenschaft und -Technologie*, vol. 149, no. 4, Article ID 111832, 2021.

[25] Y. Hou, Q. Li, Q. Han et al., “MobileCrack: object classification in asphalt pavements using an adaptive lightweight deep learning,” *Journal of Transportation Engineering, Part B: Pavements*, vol. 147, no. 1, Article ID 04020092, 2021.

[26] H. Park, “Edge based lightweight Authentication architecture using deep learning for vehicular networks,” *Journal of Internet Technology*, vol. 23, no. 1, pp. 193–200, 2022.

[27] Z. Cao, W. T. Shih, J. Guo, C. K. Wen, and S. Jin, “Lightweight convolutional neural networks for CSI feedback in massive MIMO,” *IEEE Communications Letters*, vol. 25, no. 8, pp. 2624–2628, 2021.