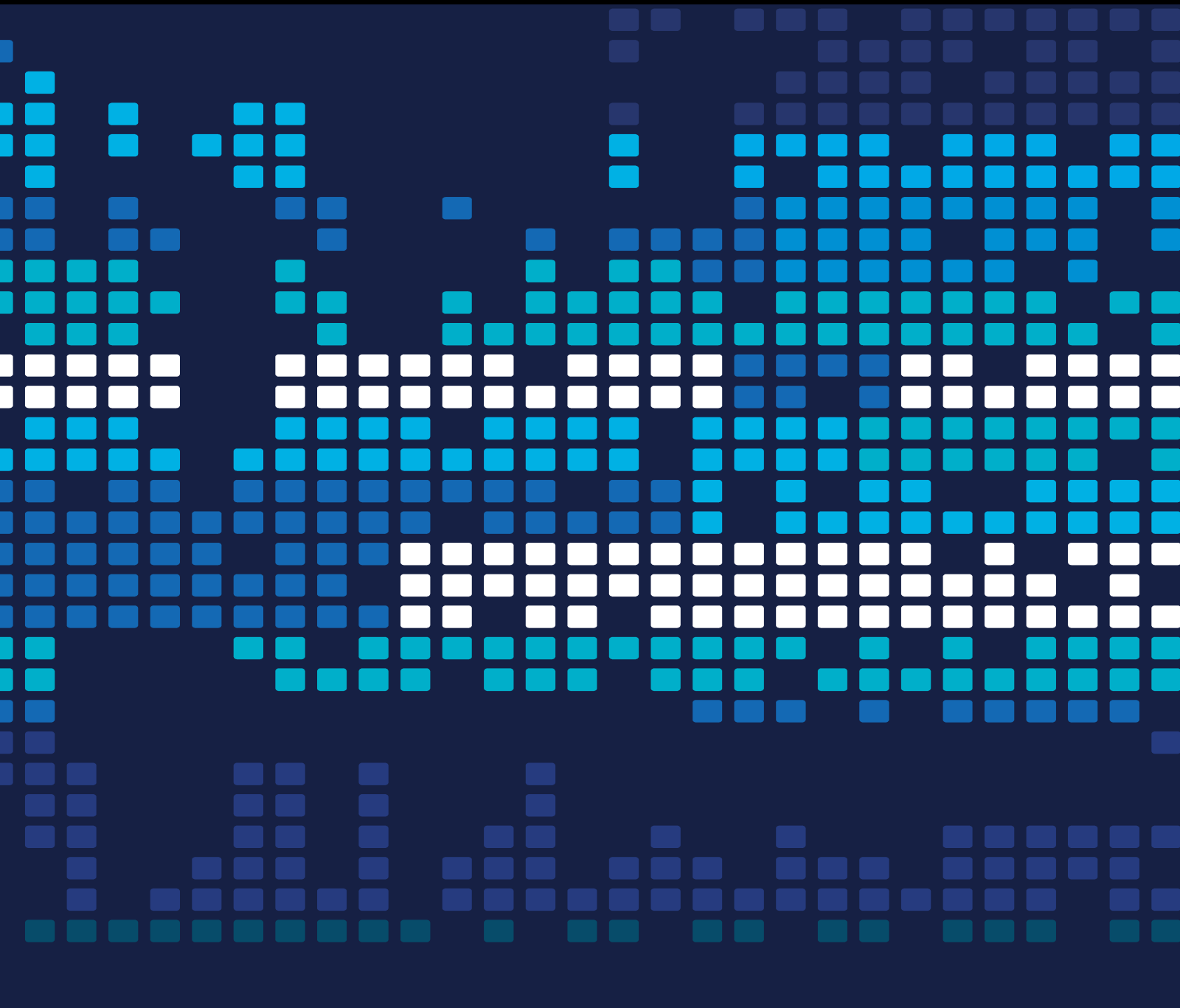


Scientific Programming Towards a Smart World 2021

Lead Guest Editor: Wenbing Zhao

Guest Editors: Chenxi Huang and Yizhang Jiang





Scientific Programming Towards a Smart World 2021

Scientific Programming

**Scientific Programming Towards a
Smart World 2021**

Lead Guest Editor: Wenbing Zhao


Guest Editors: Chenxi Huang and Yizhang Jiang



Copyright © 2022 Hindawi Limited. All rights reserved.

This is a special issue published in "Scientific Programming." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chief Editor

Emiliano Tramontana , Italy

Academic Editors

Marco Aldinucci , Italy
Daniela Briola, Italy
Debo Cheng , Australia
Ferruccio Damiani , Italy
Sergio Di Martino , Italy
Sheng Du , China
Basilio B. Fraguela , Spain
Jianping Gou , China
Jiwei Huang , China
Sadiq Hussain , India
Shujuan Jiang , China
Oscar Karnalim, Indonesia
José E. Labra, Spain
Maurizio Leotta , Italy
Zhihan Liu , China
Piotr Luszczek, USA
Tomàs Margalef , Spain
Cristian Mateos , Argentina
Zahid Mehmood , Pakistan
Roberto Natella , Italy
Diego Oliva, Mexico
Antonio J. Peña , Spain
Danilo Pianini , Italy
Jiangbo Qian , China
David Ruano-Ordás , Spain
Željko Stević , Bosnia and Herzegovina
Kangkang Sun , China
Zhiri Tang , Hong Kong
Autilia Vitiello , Italy
Pengwei Wang , China
Jan Weglarz, Poland
Hong Wenxing , China
Dongpo Xu , China
Tolga Zaman, Turkey

Contents

Using Ensemble Learning Algorithms to Predict Student Failure and Enabling Customized Educational Paths

Lassaad K. Smirani , Hanaa A. Yamani , Leila Jamel Menzli , and Jihane A. Boulahia 
Research Article (15 pages), Article ID 3805235, Volume 2022 (2022)





Research on Teaching Reform of College Student Training Mode Based on Engineering Project Economic Evaluation of Driving Behavior with Internet of Vehicles Data

Shuilong He , Zhihong Zou , Hui Li , Jucai Deng , Enyong Xu , Rongjiang Tang , and Yanyan Zhou 
Research Article (12 pages), Article ID 3805318, Volume 2022 (2022)



Enhancing Problem-Solving Ability through a Puzzle-Type Logical Thinking Game

Ting-Sheng Weng 
Research Article (9 pages), Article ID 7481798, Volume 2022 (2022)

Sizing of the Propulsion System for a Heavy-Duty Fuel Cell Commercial Vehicle

Weiguang Zheng , Weiwei Xin , Enyong Xu , Shuilong He , Jirong Qin, and Heng Wang
Research Article (16 pages), Article ID 1497178, Volume 2021 (2021)


Solving No-Wait Flow Shop Scheduling Problem Based on Discrete Wolf Pack Algorithm

Rongshen Lai , Bo Gao, and Wenguang Lin 
Research Article (6 pages), Article ID 4731012, Volume 2021 (2021)


Investor Sentiment Combined with Multisource Information to Predict Stock Prices: An Analysis of China's A-Share Market

Xin Huang  and Huilin Song 
Research Article (9 pages), Article ID 9094032, Volume 2021 (2021)

Regional Atmospheric Light Optimisation Algorithm for Heterogeneous Image Dehazing

Haoqiang Wu, Yiran Fu, Quanxing Zha, Aidong Chen, and Hongyuan Jing 
Research Article (13 pages), Article ID 3377905, Volume 2021 (2021)

Evaluation of Blended Oral English Teaching Based on the Mixed Model of SPOC and Deep Learning

Yanli Hui 
Research Article (9 pages), Article ID 7044779, Volume 2021 (2021)


Dense Dilated Attentive Network for Automatic Classification of Femur Trochanteric Fracture

Yuxiang Kang, Jie Yu, Zhipeng Ren, Guokai Zhang , Wen Cao, Yinguang Zhang , and Qiang Dong 
Research Article (7 pages), Article ID 1929800, Volume 2021 (2021)

AGVs Route Planning Based on Region-Segmentation Dynamic Programming in Smart Road Network Systems

Zheng Zhang, Juan Chen , and Qing Guo
Research Article (13 pages), Article ID 9589476, Volume 2021 (2021)

An Algorithm for Construction Project Cost Forecast Based on Particle Swarm Optimization-Guided BP Neural Network

Dan Ye 


Research Article (8 pages), Article ID 4309495, Volume 2021 (2021)

Load Forecasting Method Based on Improved Deep Learning in Cloud Computing Environment

Kai Zhang , Wei Guo , Jian Feng , and Mei Liu 





Research Article (11 pages), Article ID 3250732, Volume 2021 (2021)

Research and Design of Distributed Fire Alarm System of Indoor Internet of Things Based on LoRa

Wei Chen, ChenYu He, JianRong Lu , Kui Yan, Jin Liu, Feng Zhou, Xin Xu, and Xiao Hao


Research Article (12 pages), Article ID 7462331, Volume 2021 (2021)

Futuristic Cyber-Twin Architecture for 6G Technology to Support Internet of Everything

Sapna Juneja , Mamta Gahlan , Gaurav Dhiman , and Sandeep Kautish 



Review Article (7 pages), Article ID 9101782, Volume 2021 (2021)

Forecasting Variation Trends of Stocks via Multiscale Feature Fusion and Long Short-Term Memory Learning

Yezhen Liu , Xilong Yu , Yanhua Wu, and Shuhong Song 


Research Article (9 pages), Article ID 5113151, Volume 2021 (2021)

Material Discrimination Algorithm Based on Hyperspectral Image

Jian Zhou , Zhuping Wang, Yingjie Jiao, and Cong Nie 


Research Article (9 pages), Article ID 8329974, Volume 2021 (2021)

Impact of Check-In Data on Urban Vitality in the Macao Peninsula

Chen Pan, Junling Zhou, and Xiaohua Huang 


Research Article (9 pages), Article ID 7179965, Volume 2021 (2021)

Pedestrian Behavior Recognition Based on Improved Dual-stream Network with Differential Feature in Surveillance Video

Yonghong Tan , Xuebin Zhou, Aiwu Chen, and Songqing Zhou



Research Article (10 pages), Article ID 3279957, Volume 2021 (2021)

A Novel Brain Image Segmentation Method Using an Improved 3D U-Net Model

Zhuqing Yang 

Research Article (10 pages), Article ID 4801077, Volume 2021 (2021)




Analysis of Effectiveness and Performance Prediction of Sports Flipped Classroom Teaching Based on Neural Networks

Wei Xu , Wenying Xiong, Zhe Shao, and Yun Li 

Research Article (7 pages), Article ID 5284457, Volume 2021 (2021)



Contents

A New Embedded Estimation Model for Soil Temperature Prediction

Xuezhi Wang , Wenhui Li , and Qingliang Li 


Research Article (16 pages), Article ID 5881018, Volume 2021 (2021)

Denosing Speech Based on Deep Learning and Wavelet Decomposition

Li Wang, Weiguang Zheng , Xiaojun Ma, and Shiming Lin 

Research Article (10 pages), Article ID 8677043, Volume 2021 (2021)

A Study on the Relationship between Public Derivative Big Data and Industrial Policymaking: Taking Bike Sharing as an Example

Huilin Song 









Research Article (11 pages), Article ID 9347432, Volume 2021 (2021)

An Autocontouring Method for the Kidneys Using an Adaptive Weighted Multikernel Support Vector Machines

Yi Gu  and Bo Li


Research Article (13 pages), Article ID 1885512, Volume 2021 (2021)

Stock Price Forecast Based on CNN-BiLSTM-ECA Model

Yu Chen , Ruixin Fang , Ting Liang , Zongyu Sha , Shicheng Li , Yugen Yi , Wei Zhou , and Huilin Song 


Research Article (20 pages), Article ID 2446543, Volume 2021 (2021)

Cluster Analysis and Visualization for the Legend of the Condor Heroes Based on Social Network

Chao Fan , Zhihui Yang, and Yuyi Yuan


Research Article (10 pages), Article ID 9439583, Volume 2021 (2021)

Human-Machine Cooperation and Path Planning for Complex Road Conditions

Guanghong Zhou 

Research Article (11 pages), Article ID 7262281, Volume 2021 (2021)

Adaptive Adjustment Object Detection Algorithm under Multiple Mechanisms Based on GAN

Zemin Qiu , Feng Wang , and Zhihong Pan







Research Article (7 pages), Article ID 5875320, Volume 2021 (2021)

Applied Research of Knowledge in the Field of Artificial Intelligence in the Intelligent Retrieval of Teaching Resources

XuJing Bai  and JiaJun Li




Research Article (11 pages), Article ID 9924435, Volume 2021 (2021)

A Temporal Pool Learning Algorithm Based on Location Awareness

Lei Li , Yuquan Zhu , Tao Cai , Dejiao Niu , Huaji Shi , and Tingting Zou 

Research Article (12 pages), Article ID 9956244, Volume 2021 (2021)

Segmentation Technology of Nucleus Image Based on U-Net Network

Jie Fang , QingBiao Zhou , and Shuxia Wang 


Research Article (10 pages), Article ID 1892497, Volume 2021 (2021)

Decisions on the Orderliness and Collaborative Operation Mechanism of Each Subsystem under the Organizational Model of the Internet of Things

Lingyi Cai  and Wei Liu 



Research Article (12 pages), Article ID 2179343, Volume 2021 (2021)

Stepped Frequency Multiresolution Digital Signal Processing

Qunying Chen 

Research Article (13 pages), Article ID 9081988, Volume 2021 (2021)

Research on the Key Technologies of Network Security-Oriented Situation Prediction

Yikun Zhu  and Zhiling Du 




Research Article (10 pages), Article ID 5527746, Volume 2021 (2021)

ITDPM: An Internet Topology Dynamic Propagation Model Based on Generative Adversarial Learning

Hangyu Hu , Xuemeng Zhai , Gaolei Fei, and Guangmin Hu

Research Article (9 pages), Article ID 2390466, Volume 2021 (2021)

A Method of Amino Acid Terahertz Spectrum Recognition Based on the Convolutional Neural Network and Bidirectional Gated Recurrent Network Model

Tao Li , Yuanyuan Xu , Jiliang Luo, Jianan He, and Shiming Lin 


Research Article (7 pages), Article ID 2097257, Volume 2021 (2021)

A Medical Image Classification Model Based on Adversarial Lesion Enhancement

Bing Zhang  and Xu Hu



Research Article (9 pages), Article ID 4265650, Volume 2021 (2021)

Nonlinear Load Harmonic Prediction Method Based on Power Distribution Internet of Things

Yongle Dong, Fan Zhang, Xuan Li, Lifang Zhang, Jia Yu, Yongmei Mao, and Guanglong Jiang 

Research Article (12 pages), Article ID 9978900, Volume 2021 (2021)

An Efficient Early Frame Breaking Strategy for RFID Tag Identification in Large-Scale Industrial Internet of Things

Zhiyong He  and Hanguang Luo 

Research Article (6 pages), Article ID 4603629, Volume 2021 (2021)




The Impact of Urban Sprawl and Smart City Construction on Regional Coordination

Liangfeng Hao , Xue Chen , and Chengpeng Min 

Research Article (12 pages), Article ID 5589571, Volume 2021 (2021)

Contents

Hybrid Algorithm Based on Content and Collaborative Filtering in Recommendation System Optimization and Simulation

Lianhuan Li , Zheng Zhang , and Shaoda Zhang 



Research Article (11 pages), Article ID 7427409, Volume 2021 (2021)

Research on Multi-Target Network Security Assessment with Attack Graph Expert System Model

Yunpeng Li  and Xi Li 




Research Article (11 pages), Article ID 9921731, Volume 2021 (2021)

Deep Multiple Kernel Learning for Prediction of MicroRNA Precursors

Hengyue Shi , Dong Wang, Peng Wu , Yi Cao, and Yuehui Chen







Research Article (9 pages), Article ID 9969282, Volume 2021 (2021)

A Novel Transfer Enhanced α -Expansion Move Learning Model for EEG Signals

Jiangwei Cai , Lu Zhao , and Anqi Bi 




Research Article (12 pages), Article ID 9957845, Volume 2021 (2021)

Automatic Retinal Vessel Segmentation Based on an Improved U-Net Approach

Zihe Huang , Ying Fang , He Huang , Xiaomei Xu , Jiwei Wang , and Xiaobo Lai 









Research Article (15 pages), Article ID 5520407, Volume 2021 (2021)

Application of Internet of Things Compressed Sensing and Information Interaction Technology in Intelligent Transportation Layout

Na Li , Ze Wu , and Zhongbiao Zhao 




Research Article (11 pages), Article ID 9979198, Volume 2021 (2021)

Tumor Grade and Overall Survival Prediction of Gliomas Using Radiomics

Jianming Ye , He Huang , Weiwei Jiang , Xiaomei Xu , Chun Xie , Bo Lu , Xiangcai Wang , and Xiaobo Lai 


Research Article (11 pages), Article ID 9913466, Volume 2021 (2021)

A Distributed Congestion Control Strategy Using Harmonic Search Algorithm in Internet of Vehicles

Meiyu Pang , Jianing Shen , and Lixiu Wu 

Research Article (9 pages), Article ID 5519492, Volume 2021 (2021)

A Single-Label Model to Ensure Data Consistency in Information Security

Cigdem Bakir 

Research Article (8 pages), Article ID 9913645, Volume 2021 (2021)

Research Article

Using Ensemble Learning Algorithms to Predict Student Failure and Enabling Customized Educational Paths

Lassaad K. Smirani ^{1,2} Hanaa A. Yamani ³ Leila Jamel Menzli ⁴
and Jihane A. Boulahia ^{2,3}

¹Deanship of eLearning & Distance Education, Umm Al-Qura University, Mecca, Saudi Arabia

²InnoVCOM Lab, University of Carthage, Tunis, Tunisia

³Information Science Department College of Computer Sciences and Information Systems, Umm Al-Qura University, Mecca, Saudi Arabia

⁴Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia

Correspondence should be addressed to Lassaad K. Smirani; lksmirani@uqu.edu.sa

Received 4 November 2021; Revised 14 December 2021; Accepted 16 March 2022; Published 14 April 2022

Academic Editor: Chenxi Huang

Copyright © 2022 Lassaad K. Smirani et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the challenges in e-learning is the customization of the learning environment to avoid learners' failures. This paper proposes a Stacked Generalization for Failure Prediction (SGFP) model to improve students' results. The SGFP model mixes three ensemble learning classifiers, namely, Light Gradient Boosting Machine (LGBM), eXtreme Gradient Boosting machine (XGB), and Random Forest (RF), using a Multilayer Perceptron (MLP). In fact, the model relies on high-quality training and testing datasets that are collected automatically from the analytic reports of the Blackboard Learning Management System (i.e., analytic for learn (A4L) and full grade center (FGC) modules). The SGFP algorithm was validated using heterogeneous data reflecting students' interactivity degrees, educational performance, and skills. The main output of SGFP is a classification of students into three performance-based classes (class A: above average, class B: average, class C: below average). To avoid failures, the SGFP model uses the Blackboard Adaptive Release tool to design three learning paths where students have to follow automatically according to the class they belong to. The SGFP model was compared to base classifiers (LGBM, XGB, and RF). The results show that the mean and median accuracies of SGFP are higher. Moreover, it correctly identified students' classifications with a sensitivity average of 97.3% and a precision average of 97.2%. Furthermore, SGFP had the highest F1-score of 97.1%. In addition, the used meta-classifier MLP has more accuracy than other Artificial Neural Network (ANN) algorithms, with an average of 97.3%. Once learned, tested, and validated, SGFP was applied to students before the end of the first semester of the 2020-2021 academic year at the College of Computer Sciences at Umm al-Qura University. The findings showed a significant increase in student success rates (98.86%). The drop rate declines from 12% to 1.14% for students in class C, for whom more customized assessment steps and materials are provided. SGFP outcomes may be beneficial for higher educational institutions within fully online or blended learning schemas to reduce the failure rate and improve the performance of program curriculum outcomes, especially in pandemic situations.

1. Introduction

Today, the world is experiencing an unprecedented acceleration in the changing of life forms owing to the development of information and communication technologies (ICT). ICT continuously affects the social, economic, and cultural aspects of life. Many countries are trying to achieve success in devoting the information society [1] through

infrastructure and policies to boost knowledge acquisition and become smart societies. In the academic domain, and owing to the rapid development of technological networks and smart devices, many Learning Management System (LMS) solutions have emerged [2].

During the Covid-19 pandemic, orientation toward e-learning and distance education has become necessary, and academic users are being more convinced about the

importance of these new learning trends. In fact, universities worldwide have shifted from face to face-teaching mode to fully online learning with lockdown measurements imposed by governments and the need for social distancing. Thus, most educational institutions seek to rely on fully online or blended learning models to ensure educational process continuity [3]. Besides, the increasing use rate of the new learning and teaching models, the results show that many students are not familiar with this new environment and are stressed [4]. In fact, some studies mentioned that [5–9], and [10], increasing failure and dropout rates are observed. This is mainly due to unsuitable assessments and materials for students with learning and teaching difficulties. Today, the efficiency of online learning and teaching has become one of the most important research areas discussed in the educational community.

In this context, many researchers have concentrated their efforts on developing new solutions to address these challenges. The first researchers' wave is oriented towards adaptive learning mode and recommended systems as solutions for these challenges, such as [7–12] and [13]. They state that adaptive education systems can significantly support students' achievements via customized environments (i.e., teaching processes, materials, and assessments). Nevertheless, the quality of teaching processes and subject materials remains a significant challenge for these systems [14–17].

The second wave of researchers attempted to find solutions for predicting students' failures. In fact, the rise of Artificial Intelligence (AI) has facilitated the development of a series of predictive models based on electronic online assessment and LMS tools. Faculty members are free to intervene and avoid student failures during learning, teaching, and assessment processes [5, 18–34].

Most of these studies use at most one predictive technique and have problems with performances and especially with low accuracy. Ensemble learning techniques are used as a solution for this issue, but it is necessary to consider the compromise between the number of systems used, the complexity, and the desired outcomes. Ensemble learning increases global accuracy by fusing the predictions of many learners. The most well-known learning techniques are bagging, boosting, and stacking. In this study, the ability of stacked generalization for classification is used to ensure customized teaching and learning tasks. The classification parameters can be the learner's strengths, skills, interests, and needs. Each student's class has a specific learning path to overcome learning weaknesses and guarantee success [27, 35–37].

In this study, a Stacked Generalization for Failure Prediction (SGFP) is used to enhance the classification performance of students and the prediction of their results. The proposed system allows dropout prediction and intervening in offering different learning paths to guarantee students' success and reduce failure. Data is extracted from the Blackboard Learning Management System and each learning path is designed according to the students' classes by the "Adaptive Release" Blackboard tool.

The organization of this paper is as follows: Section 2 presents the literature review where studies using ensemble learning techniques for prediction are presented and discussed. Section 3 explains the methodology of the study. Results analyses are detailed in Section 4, where we discuss the performance of our proposed approach. Section 5 describes the conclusion concerning this proposed system.

2. Literature Review

The effectiveness of online education is being debated once more, particularly during the pandemic era. Today's educational institutions are expected to develop new educational solutions that are not bound by time or space. They need to create attractive learning environments and strategies to meet these expectations through e-learning, blended learning, mobile learning, and online education. Faculty members are also encouraged to create activities that allow students to actively explore and build their understanding of a given topic. They should succeed in their distance-learning courses. Furthermore, they must plan activities and learning tasks in a flexible manner while providing feedback to allow students to progress at their own accelerated rates [38–40]. E-learning systems concepts and recommendation methodology in the educational adaptation context have been the focus of several studies [7,8, 41].

In order to improve the quality of distance education, two main research areas of interest through LMS are identified: the first is concerned with adaptive learning and recommended systems, and the second is concerned with prediction of students' performances.

2.1. Adaptive Learning and Recommended Systems. The first research line attempts to better understand students' knowledge levels, as well as to recognize their preferences and how to learn and understand specific concepts based on students' educational interests. References [9,15,42, 43].

In [10,11,44, 45], e-learning platforms are oriented to attract educational users and place electronic materials according to their preferences.

Shekapure and Patil [12] adopted a personalizing e-learning method that provides variable learning objects for students. The proposed method uses customized data, such as student knowledge and student learning style, to customize learning paths according to students' profile classification. Indeed, the proposed approach is cognitive, progressive, and dynamic, where instructors can align e-learning material and sequencing to learners' profiles and performances.

Cerna et al. [13] designed an innovative system that aims to impose an order in managing electronic course content accompanying topics related to the geography, history, and culture of English-speaking countries. Their proposed approach is a blended learning model that uses LMS tools for communication, navigation, and evaluation of student presentations.

Keskin et al. [41] discussed the individual factors represented in cognitive learning strategies: the extent of

readiness for e-learning, and the motivation factor. The results of their study showed a correlation between learning environment preferences and the self-efficacy factor between e-learning motivation and task value, between the learning environment and self-efficacy constructs, as well as between e-learning motivation and task value. However, cognitive strategies, self-directed learning, student control, and anxiety factor testing were independent of students' preferences for lecturing.

2.2. Prediction of Students' Performances in E-Learning Environment. During the last few years, students' performance assessment and improvement has been an important objective for all higher institution parties. This is in harmony with the workforce requirements of highly skilled and competent employees. Faculty members are expected to deploy all strategies, methods, and tools to motivate students to enhance their skills and competencies, and thus their grades, either in face-to-face or distance-learning environments. A new Educational Data Mining (EDM) focuses on predicting students' performances in e-learning environments [46–49]. EDM applies supervised or unsupervised machine learning algorithms to inspect, analyze, and learn educational data, and then predict students' performance.

The second research line introduced intelligent methods. Predicting students' performance using statistical and machine learning techniques is no novel [43,50]. Although these works have been fruitful, applying the Artificial Neural Network (ANN) model is still broad for e-learning students' performances compared to its use in other domains [51]. ANNs are imitations of human brain neuron functions to solve machine learning problems, and their use essentially leads to intelligent behavior. ANNs can be used to solve many types of problems such as forms, speech recognition, and function approximation, but their application in the eLearning field is especially based on classification and prediction [52–54].

Arsad et al. [55] presented an ANN-based system for early performance prediction. The study was conducted with engineering students at a Malaysian university. In fact, academic achievement in semester eight was measured using the cumulative grade point average. Moreover, Adewale et al. [56] used a feedforward neural network applied to secondary school students to study the relationship between cognitive and psychological variables on academic performance. The authors concluded that clustering students according to their performance into different categories using ANN is an efficient method that enables curriculum developers and educational planners to provide better educational services. In [21], the authors utilized data mining to identify patterns and student grouping used to explain academic dropout. They collected data from students who registered for two admission periods at the Universidad Tecnológica Indoamerica in Ambato, Ecuador. They classified and defined the performance patterns using a k-means algorithm, and predictions for new students were made using a Support-Vector Machine (SVM) model. The study in [24] predicts grades and proposes a deep learning model

comprising scattered attention layers, convolutional neural layers, and a fully connected layer. Grades, student demographics, and course descriptions were among the information gathered. The proposed model achieved 81% prediction accuracy and 85% failure prediction accuracy and provided a potential explanation for the predicted outcome. In [6], the study employs machine learning models to improve the prediction of previous student performance and explain why a student's performance reaches a certain score. It also provides a visual technique to assist in determining the factors that most influence the score through the experiment, allowing educators to identify students at-risk early and provide appropriate exhortation in an advantageous way.

Furthermore, Kalyani et al. [36] applied a convolutional neural network (CNN) model to predict and assess student performance. The number of hours the student spent studying and the degree of student involvement in academic activities were used as predictor variables. In [37], the recurrent neural network (RNN) model was used to predict the final grade and was compared to a multiple regression analysis model, in which RNN was applied for early prediction of 108 students' results and confirmed its effectiveness.

A comparison study was conducted in [57] to predict students' academic performance based on a single performance factor. Both multilayer perceptron neural network (MPNN) and generalized regression neural network (GRNN) learning algorithms are employed on collected data from documents and student transcript records. The findings *t* reveal that the overall performance of GRNN outperforms the MPNN Multilayer Perceptron model with an accuracy of 95%. In addition, the study concluded that GRNN could be used by educators to predict student academic performance based on a single performance factor.

Although the mentioned models provide a high capacity for classification and prediction of student failures, they rely on individual classifiers that have little knowledge of the dataset. Hence, the use of the ensemble learning method increased the average prediction performance.

2.3. Ensemble Learning Approach for Prediction in E-Learning Environment. The stacked generalization model is one of the most well-known learning techniques. It combines many learners and uses their outcomes as an input to the meta-learner to predict the final student class. Previous studies have deployed the stacked model in various fields to increase the accuracy prediction and decrease the lowest prediction error. For instance, in the education field, the authors in [35] presented a stacked generalization model composed of three learners: back propagation neural networks, support vector machines, and M5P model tree. This study aims to predict academic achievement after graduating from students. The authors used the root mean square error (RMSE) to evaluate the model performance. The prediction result of stacking compared to the three classifiers was better. In [58], the authors presented a model that combines stacking and voting ensembles to evaluate faculty performance. They used

two datasets: the first is from the UCI machine learning repository called the Teaching Assistant Evaluation; hence, the second is from the students of a university. They employed 15 algorithms. When compared to methods using only one model, the proposed method produced higher accuracy, but the problem of complexity arises because the use of 15 algorithms generates a delay in execution; hence, there is a need to respect the compromise between the desired results and the complexity.

Alizamar et al. [59] presented a stacked RASCH model to represent the differences in aggressive behavior between male and female students at the Junior High School of West Sumatera.

The authors in [60] used ensemble learning to identify style learning based on the Vark model (visual, auditory, reading/writing, and kinesthetic). They have used J48, SVM, Random Forest, Naive Bayes, and hard majority voting. The results obtained in the 20% Test set are as follows: the J48 decision tree achieved 57.2% precision level and 61.4% accuracy, the SVM classifiers achieved 58.5% precision level and 59.6% accuracy. The Random Forest achieved 57.9% precision level and 60.5% accuracy. The Naïve Bayes achieved 59.7% precision level and 62.3% accuracy. The Hard Majority voting achieved 60.2% precision level and 62.9% accuracy.

In [63], the authors evaluated eight classification techniques in order to recognize the parameters that make a significant contribution to providing an excellent model for classifying a student based on his performance. The J48 Decision Tree classifier achieved 93.5% precision, 93.5% recall and 93.2% F1-score. The Logistic Regression achieved 91% precision, 89.6% recall and 90.3% F1-score. The Multi-Layer Perceptron achieved 92.5% precision, 90.5% recall and 91.2% F1-score. The Support Vector Machine achieved 96% precision, 89% recall and 92.4% F1-score. The AdaBoost (Adaptive Boosting classifier) achieved 90% precision, 85.9% recall and 92.4% F1-score. The Bagging achieved 96.9% precision, 92.4% recall and 91.8% F1-score. The Random Forest achieved 97% precision, 90.8% recall and 93.8% F1-score. The Voting achieved 93.1% precision, 91.4% recall and 92.3% F1-score.

In [64], the authors predicted student academic performance by proposing a Hybrid Ensemble Learning Algorithm (HELA). The Super Learner algorithm receives prediction results from base classifiers such as Gradient Boosting, Extreme Gradient Boosting, Light Gradient Boosting Machine, and various combinations of these algorithms. A Random Search algorithm is used to optimize the hyper-parameters of base classifiers. The proposed algorithm predicts students' performance in two courses, and the experimental results achieved 96.6% and 91.2% accuracy values.

The study in [65] suggested an ensemble learning method based on label distribution estimation called light gradient boosting channel attention network (LGBCAN). This model is employed to forecast performance in web-based learning tasks. The Channel Attention Network (CAN) model enhances LightGBM's function by concentrating on better outcomes in the K-fold cross entropy of

LightGBM. The LGB achieved 68.33% accuracy, 68.33% precision, 59.03% recall and 61.58% F1-score. The XGB achieved 67.71% accuracy, 62.45% precision, 58.63% recall and 60.47% F1-score. The LGBALD achieved 57.84% accuracy, 63.73% precision, 57.14% recall and 60.26% F1 score. The XGBALD achieved 56.76% accuracy, 63.23% precision, 56.96% recall and 59.93% F1-score. The LGBCAN achieved 68.14% accuracy, 63.56% precision, 63.66% recall and 63.61% F1-score.

Likewise, most of the mentioned studies use datasets collected from questionnaires, surveys, student registration units, and students' transcripts to train and test models [27, 36, 37, 55–58, 61].

The weakness of some studies is related to the potential issue with data quality, which is in some way outdated, inaccurate, subjective, and does not reflect the real students' activities through e-learning processes. It is obvious that to better predict students' failure and dropout, significant and objective data will be of great help. Indeed, learning algorithms/models' results in e-learning environments are more interesting if data are directly used from LMS platforms.

To deal with these limitations, this study presents the idea of exploiting LMS solution analytical reports. In fact, all LMS platforms provide an amalgam of easy-to-use and complete analytical generating report tools. These reports are interesting sources of knowledge to investigate: (i) all students' assessment grades are there, (ii) the content reflects the real students' e-learning activities (i.e., access to course contents, assignment submissions, interactions, etc.), and (iii) the data is clean and well structured. However, to the best of our knowledge, this perception is still lacking.

This study focuses on offering the SGFP model to avoid student failures by exploiting LMS analytic reports for a specific set of courses and for a fixed academic year. The study was conducted to predict students' performance and avoid dropout. The training and testing datasets are extracted automatically from Blackboard analytic reports through its Analytic for Learn (A4L) and Full Grade Center (FGC). The SGFP algorithm is based on heterogeneous data reflecting students' interactivity degrees on one side and students' educational performances and skills on the other. These data are more significant and better for the e-learning process and evaluation. Extracting updated data from real student activities offers accuracy, objectivity, and precision. SGFP uses the Blackboard adaptive release tool to customize learning paths for each student's class.

3. Methodology

The research methodology of this study was based on the classification capacities of the stacked generalization. It is composed of six steps: (1) data collection, (2) data pre-processing, (3) SGFP modeling, (4) training and testing, (5) evaluation, and (6) validation.

3.1. Data Collection. The LMS creates a large amount of data. Universities are optimizing instructional design, increasing faculty effectiveness, and identifying at-risk students in time

to keep them on track to graduate with high-quality credentials by deliberately presenting this data in the form of information. Like any LMS, Blackboard provides educators with a better way to access the information they need to assist and maintain their students' progress. Instructors can access course dashboards and other reports from Blackboard A4L directly from Blackboard Learn courses. All students enrolled in the class are listed in Table 1: T: theoretical course; P: practical course).

The data extracted included 376 students enrolled from the start of the 2019-2020 academic year to the end of the first semester of the 2020-2021 academic year. Figure 1 shows SGFP training and testing data extracted from Blackboard for 22 courses from the Information Sciences Department (ISD) at the College of Computer and Information System (CCIS) and the Computer Science Department (CSD) at the First Common Year College (FCYC).

3.1.1. Input Data. Table 2 lists the input and output of the SGFP dataset. The dataset was collected via Blackboard tools as follows:

A4L: This allows instructors to keep track of how students are performing through running course reports. A4L extracts data from the UQU system information science (SIS) and Blackboard-Learn. Four types of reports are provided from A4L: Course at-a-Glance, activity grade scatter plot, activity matrix, and course submission summary report. The SGFP classification is based on six predictor variables: five variables from the Blackboard-Course at-a-Glance report:

- (i) Access Operations (AO): this indicator refers to the number of operations performed by the student.
- (ii) LMS time spent (LTS): this metric indicates how long a student spends per week browsing the online course on Blackboard.
- (iii) Degree of student interactions (DSI): this indicator measures during the week, the degree of

Student interaction with the course and its tools, interactions with the teacher, and interactions with his colleagues.

- (i) School Assignment Submission (SAS): this indicator reflects students' commitment to homework and assessment tools.
- (ii) Grade of Evaluation Center (GEC): this parameter represents the grade obtained by students during subsequent participation through Blackboard.

FGC: this is a container for all students' assessment grades, and the sixth input variable, Grade of Initial Exercise (GIE), was extracted from FGC.

The SGFP model entry matrix is composed of seven columns (six columns for input and one column for output) and 376 rows. Equal low weights (12%) were assigned to the five first inputs, and the highest weight (40%) was assigned to the sixth variable (GIE). On the one hand, the rationale for the weights associated with these parameters is that students' activities on the LMS platform cannot always be meaningful.

Some students can hardly navigate the LMS, download course content, work seriously offline, and achieve good results. Therefore, the highest weight is for the initial exercise evaluation as it more reflects the students' performance and is decisive to e-learning course customization. On the other hand, equitable weighting was chosen for the five A4L indicators because their impact on student rankings is similar, and they represent students' interest in digital education and measure their virtual interactivity with online courses, but they do not reflect their actual scientific performance. Weak weights are given because the parameters that represent the impact of browsing time, operations performed, interactions with LMS content. All have the same degree of importance in evaluating student performance.

3.1.2. Output Data. Students' profiles are classified into three classes (A, B, and C) according to the six predictor variables. For each class, a special learning path is fixed as follows:

For class A, students were allowed to take the exam directly without taking any other support courses. Their learning path offers video sequences containing the essential summaries of the courses.

For class B, students with grade above average might present a risk of not obtaining 60% in the final exam. The learning path offered these students descriptive summaries and short-corrected tests covering all course modules.

For class C, to avoid dropout, a customized learning path is then offered by summing up the descriptive summaries and recapitulative units. Each unit ends with an evaluation test that is essential for students to move from one unit to another. The serious application of this learning-path customization can allow for an excellent success rate.

The Blackboard Adaptive Release tool was used to customize learning paths. Adaptive Release is a dedicated publishing tool that allows instructors to deliver personalized content to students. This content is based on a set of rules related to four criteria: date, membership, grade, and review status. The goal of personalizing the publication of content is to create courses that are more interactive and tailored to the needs of each learner.

3.2. Data Preprocessing. SGFP learning data were extracted from the Blackboard Course at-a-Glance report and from the FGC tool one month before the final exam (Figure 2). Then, they were normalized according to the weights assigned to them. These variables will then be put in an Excel file. The columns of this file represent the five normalized variables extracted from A4L, while the rows represent the students' names.

The sixth input variable was taken from the FGC. This variable must be normalized according to the weight assigned to it, then it will be placed in the sixth column of the Excel file.

TABLE 1: The courses concerned by Dataset.

Course name	No. of students	Section	Faculty/department	Method of learning
Networks (T)	41	2	CCIS/ISD	Blended learning.
Networks (P)	41	2	CCIS/ISD	Blended learning
Website development 2 (T)	42	1	CCIS/ISD	Blended learning
Website development 2 (P)	42	1	CCIS/ISD	Blended learning
Analyse and design of information systems (T)	35	1	CCIS/ISD	Blended learning
Analyse and design of information systems (P)	35	1	CCIS/ISD	Blended learning
Analyse and design of information systems (T)	34	2	CCIS/ISD	Blended learning
Analyse and design of information systems (P)	34	2	CCIS/ISD	Blended learning
Analyse and design of information systems (T)	37	4	CCIS/ISD	Blended learning
Analyse and design of information systems (P)	37	4	CCIS/ISD	Blended learning
Programming python 1 (T)	34	1	FCYC/CSD	Fully online
Programming python 1 (P)	34	1	FCYC/CSD	Fully online
Information technology (T)	10	5	CCIS/ISD	Fully online
Information technology (P)	10	5	CCIS/ISD	Fully online
Programming python 1 (T)	14	1	CCIS/ISD	Fully online
Programming python 1 (P)	14	1	CCIS/ISD	Fully online
Networks (T)	41	4	CCIS/ISD	Fully online
Networks (P)	41	4	CCIS/ISD	Fully online
Website development 2 (T)	50	1	CCIS/ISD	Fully online
Website development 2 (P)	50	1	CCIS/ISD	Fully online
Website development 2 (T)	38	2	CCIS/ISD	Fully online
Website development 2 (P)	38	2	CCIS/ISD	Fully online
Total no. of students	376			

Grade of evaluation center		School assignments submission		Degree of student interactions		LMS time spent		Access operations		Last Assessment	Last Access
oaverage	student	oaverage	student	oaverage	student	oaverage	student	oaverage	student		
87%	→ 80%	2	→ 2	90	↓ 66	267	↓ 57	14	↓ 6	26/02/2020	27/04/2020
87%	↓ 60%	2	→ 2	90	↑ 185	267	↑ 294	14	↑ 28	26/02/2020	22/04/2020
87%	→ 80%	2	→ 2	90	↑ 104	267	↓ 186	14	↓ 9	26/02/2020	05/04/2020
87%	→ 90%	2	→ 2	90	↓ 78	267	↓ 120	14	→ 15	26/02/2020	19/04/2020
87%	→ 90%	2	→ 2	90	→ 87	267	→ 261	14	↓ 10	26/02/2020	21/05/2020
87%	→ 90%	2	→ 2	90	→ 87	267	↓ 73	14	↑ 16	26/02/2020	12/04/2020
87%	→ 90%	2	→ 2	90	↓ 59	267	↓ 47	14	↓ 3	26/02/2020	12/05/2020
87%	→ 90%	2	→ 2	90	→ 87	267	↓ 57	14	↑ 16	26/02/2020	24/04/2020
87%	↑ 100%	2	→ 2	90	↑ 118	267	↑ 859	14	↑ 20	26/02/2020	09/05/2020
87%	→ 90%	2	→ 2	90	↓ 63	267	↓ 153	14	→ 13	26/02/2020	22/04/2020
87%	↑ 100%	2	→ 2	90	↑ 106	267	↑ 553	14	↑ 17	26/02/2020	07/05/2020
87%	→ 90%	2	→ 2	90	→ 85	267	↓ 132	14	↑ 18	26/02/2020	12/04/2020
87%	→ 90%	2	→ 2	90	→ 82	267	↑ 334	14	↑ 19	26/02/2020	21/04/2020
87%	→ 80%	2	→ 2	90	↓ 63	267	↓ 83	14	↓ 6	26/02/2020	22/04/2020
87%	→ 90%	2	→ 2	90	↓ 79	267	↑ 802	14	↑ 17	26/02/2020	21/04/2020

FIGURE 1: LMS data extracted from A4L, used for SGFP inputs.

Regarding SGFP targets, the data are extracted from the Blackboard-FGC tool, and simple programming in Excel allows the organization of this variable, which is the seventh column of the Excel file. Three values were used for this response variable. The data placed in order in the Excel file were read by our proposed model to ensure the learning stage.

All data were extracted from Blackboard and computed in the same manner. Blackboard determines the highest level of activity within a given class and course, and then computes the mean for each parameter. All students in the same class will be ranked in relation to the average, with a green arrow pointing up if they are above the mean value, and a red arrow pointing down if they are below the mean value. After

assigning a weight to each parameter, we proceeded in a standard manner to normalize these variables, as

$$\text{normalized value} = \frac{\text{extracted value} * 12}{\text{maximum value}}. \quad (1)$$

3.3. SGFP Modeling. Stacked generalization or stacking is an ensemble learning technique that is used to fuse diverse machine-learning algorithms.

This is a hierarchy of learners' levels. More specifically, the stacking is contained in two levels: the first is composed of the base learners, while the second is composed of a meta-

TABLE 2: SGFP dataset.

Training variables	Abbr.	Types of data	Source
Access operations	AO	Input	A4L: Course at-a-Glance
LMS time spent	LTS	Input	A4L: Course at-a-Glance
Degree of student interactions	DSI	Input	A4L: Course at-a-Glance
School assignments submission	SAS	Input	A4L: Course at-a-Glance
Grade of evaluation center	GEC	Input	A4L: Course at-a-Glance
Grade of initial exercise	GIE	Input	Full grade center
Student class	SC	Output	Full grade center

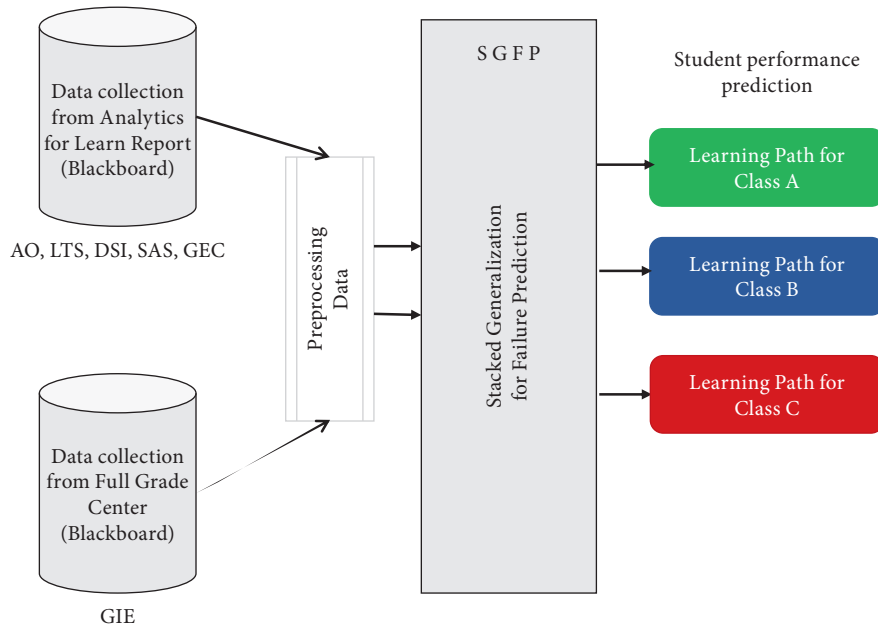


FIGURE 2: Conceptual framework for Gathering Data for SGFP learning stage.

learner. The main idea behind the stacked generalization, besides the combination of the prediction values from the base learners, is to improve the prediction performance over all the learners in the ensemble. The stacked method was employed for failure prediction. In this work, four different learners named Random Forest (RF), XGBoost (XGB), Light Gradient Boosted Machine (LGBM), and Multilayer Perceptron (MLP), were used to form a stacking model. All the first-level classifiers of our model are ensemble learning models based on decision tree algorithms.

3.3.1. Base Learners

- (i) Random Forest (RF) is an ensemble method built on decision trees. Indeed, it comprises many trees. Each tree in the forest provides a classification, known as tree votes. The forest fusions all vote trees and chooses the highest vote as the final prediction. Random forest is an efficient method and is highly accurate because of the number of decision trees used in the process.
- (ii) Extreme Gradient Boosting (XGB) is based on the principles of the gradient-boosting framework. It is

a tree ensemble approach in which the trees are added sequentially, and each tree learns from its predecessors, where they aim to minimize the errors of the previous tree. The trees are provided in parallel tree boosting to solve tasks quickly and accurately. XGboost is created to thrust the extreme computational limits of machines to provide a scalable and efficient library.

- (iii) Light Gradient Boosted Machine (LGBM) is a decision tree ensemble method for regression and classification tasks. It divides the tree leafwise with the best fit. LGBM reduces the loss of the level-wise algorithm when it is growing on the same leaf, which makes the predictions more accurate. LGBM is a fast, efficient, and distributed gradient-boosting framework.

3.3.2. Metalearner

- (i) *Multilayer Perceptron (MLP)* is a type of feedforward artificial neural network. It is inspired by the sophisticated functionality of human brains, where there are hundreds of billions of interconnected neurons. The MLP includes three layers: the input

TABLE 3: Parameter settings of the sgfp model.

Machine learning model	Parameter	Value
LGBM	Number of estimators	100
	Learning rate	0.1
	Max depth	-1
XGB	Number of estimators	100
	Learning rate	0.3
	Max depth	6
RF	Number of estimators	200
	Learning rate	0.2
MLP	Max depth	3
	Max iterations	1000
	Hidden layer sizes	200

layer, hidden layer, and output layer. MLP works using the backpropagation technique for trains.

3.4. SGFP Parameters. The specific parameter settings of different machine learning algorithms are presented in Table 3.

3.5. SGFP Steps. The algorithm below describe the process of the SGFP model in this work:

The framework of the proposed method is shown in Figure 3. The groups of students were divided for each section into two subgroups: the first group represented 90% of the students reserved for learning, and the second group was reserved for testing the model. The K-fold cross-validation technique is used to avoid the use of the same training dataset at both method levels, particularly to avoid overfitting.

In the first step, the divided dataset was used to train and test the first-level base learners of stacking. In the next step, the meta-dataset is generated from the predictions of the first-level models. Then, the MLP is trained and tested using the constructed feature instances to produce the student class. Students are classified into three groups according to their final marks, as shown in Table 4 and Figure 4.

- (i) Class A: student's grade between 80 and 100
- (ii) Class B: student grades between 60 and 80.
- (iii) Class C: student grades less than 60.
- (iv) The input vectors and target vectors were randomly divided into two sets as follows:
- (v) 90% were used for training.
- (vi) The remaining 10% was used as a completely independent test of the SGFP model.

3.6. SGFP Evaluation. To evaluate our proposed SGFP model at different levels, it was compared with the employed base classifiers in terms of accuracy, precision, and recall. For given input data, a binary classifier generates output with two class values 1/0. The one of interest is typically represented as "positive," while the other is denoted as "negative." The observed labels for all data

instances are contained in a test dataset used for performance evaluation. Having followed classification, the observed labels are compared to the predicted labels to determine performance. If a binary classifier's performance is perfect, the predicted labels will be identical, but it is relatively rare to be able to develop an ideal binary classifier that is useful in a variety of situations.

The confusion matrix is constructed from the three components of binary classification. A binary classifier predicts whether all data instances in a test dataset are positively or negatively. True positive, true negative, false positive, and false negative are the four outcomes of this classification.

- (i) True positive (TP): correct positive prediction
- (ii) False positive (FP): incorrect positive prediction
- (iii) True negative (TN): correct negative prediction
- (iv) False negative (FN): incorrect negative prediction

Accuracy is the percentage of correct predictions that a learner has achieved. It is computed by dividing the number of correct estimates by the total number of predictions:

$$accuracy = \frac{TN + TP}{TN + FP + FN + TP} \quad (2)$$

- (i) Precision, also known as the positive predictive value, is the ratio of the pertinent instances to the retrieved instances:

$$precision = \frac{TP}{FP + TP} \quad (3)$$

- (i) Recall, also called sensitivity, is a fragment of the retrieved relevant instances:

$$recall = \frac{TP}{FN + TP} \quad (4)$$

- (i) The F1-score is a statistical measure that combines precision and recall with rate performance.

$$F1 - score = \frac{precision * recall}{precision + recall} \quad (5)$$

As shown in Figure 5 and Table 5, the accuracies of all the base classifiers excluding the LGBM model vary when dealing with three test samples resulting from the 10-fold cross-validation technique.

Although they are some classification problems, it is obvious that some methods perform better in one and not so well in the other. However, the SGFP method exhibited consistent and high accuracy. Moreover, the mean and median accuracies for the SGFP model are slightly higher than those of the other classifiers.

To demonstrate the advantage of the choice of the MLP model as a meta-classifier, it was compared with other models in terms of accuracy. It is obvious that the MLP algorithm outperforms the other algorithms, with an average accuracy of 97.3%.


```

Input: Data sets:  $D = \{X_i, Y_i\}_{i=1}^m$ 
Where  $X_i$  the features  $\{X_i \in \text{AO,ITS,DSI,SAS,GEC, GIE}\}$ 
and  $Y_i$  is the labels  $\{Y_i \in \text{A,B,C}\}$ .
List of learners  $L_n$ 
Where the three base learners are:  $L_1 = \text{RF}$ 
 $L_2 = \text{XGB}$ 
 $L_3 = \text{LGBM}$ 
and the meta-learner is:  $L = \text{MLP}$ .
Output: the class label.
Step 1: learn the first-level models (Base Learners): for  $k=1$  to  $n$  do
  Train  $L_k$  based on  $D$ 
end for
Step 2: create the new data sets  $D'$  from the output of the first level:
  for  $i=1$  to  $m$  do
    Create a new data sets  $D' = \{X_i, Y_i\}$ , where
     $X_i = \{C_1(X_i), C_2(X_i), C_3(X_i)\}$ 
  end for
Step 3: learn the Second level (metalearner): train the meta-learner  $L$  based on  $D'$ 
  Make the final prediction
  Return Class labels
    
```

ALGORITHM 1: SGFP Algorithm.

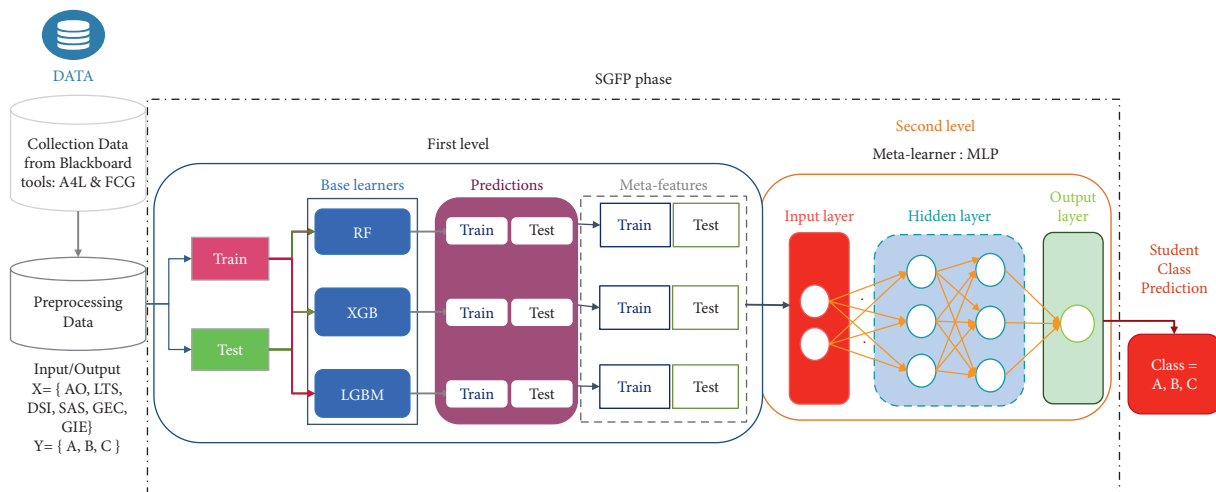


FIGURE 3: Sgfp flowchart.

Table 6 shows that our model successfully recognized the three student classes with a sensitivity average of 97.3% and that the prediction is valid when the relevant class is estimated with a precision average of 97.2%. Moreover, our model had the highest F1-score (97.1%).

This was expected, given that the proposed technique is built on a combination of an ensemble of failure prediction models rather than predictions of single learners (LGBM, XGB, and RF).

4. SGFP Validation

Having learned and tested, the proposed model can classify students into three classes. The validation step of the SGFP was carried out on 176 students who were studied during the academic year 2020-2021. The data retrieved from Blackboard were for the courses listed in Table 7.

The ranking results will be taken into consideration to guide students towards educational paths. To guarantee this mission, Blackboard’s “Adaptive Release” tool is used to enforce the follow-up of the paths traced for the students. After training the SGFP model, the program was ready to automatically classify every new student.

The student thus classified in a very specific group would be obliged to follow an educational path before taking the final exam. Blackboard provides a dedicated publishing tool called “Adaptive Release” that allows instructors to deliver personalized content to students. This content is based on a set of rules created by the instructor. The rules are related to four criteria: date, membership, grade, and review status.

Figure 6 depicts three flowcharts that correspond to three classes of students. Each student will be able to automatically follow a learning path and this according to the class he is belonging to.

TABLE 4: Students' classes.

Group	Grades	Number of students
A	$80 \leq \text{grades} < 100$	250
B	$60 \leq \text{grades} < 80$	106
C	Grades < 60	20

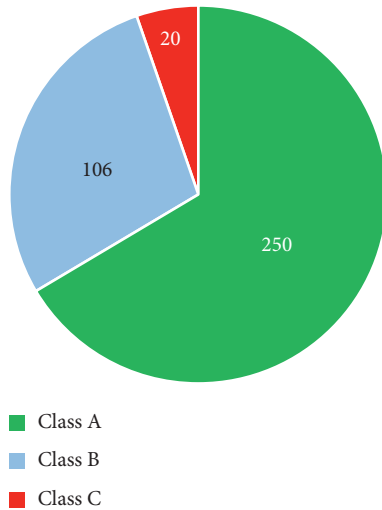


FIGURE 4: Students classification groups.

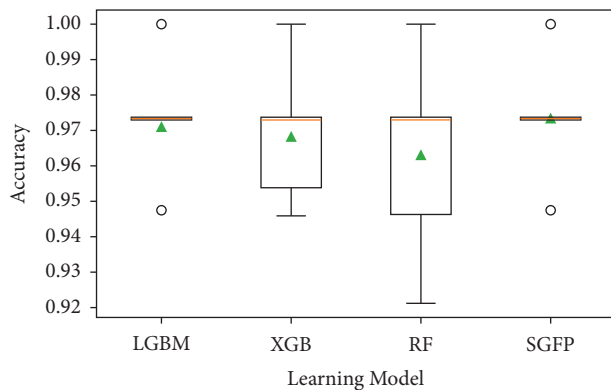


FIGURE 5: Box-plot distributions of classification accuracy for the 3 test samples resulting from the 10-fold cross-validation technique.

The goal of personalizing the publication of content is also to create courses that are more interactive and tailored to the needs of each student. Blackboard offers two types of adaptive versions: basic and advanced. For the first type, instructors apply only one rule to a piece of content for all criteria types. Students are required to meet all the rules' criteria before the item (e.g., file, image, video) is published. Students belonging to class C had to follow a complicated path with specific criteria and item output restrictions. It is obvious that the path course for class C students is more difficult to access. The second type of "Adaptive Release" is advanced, and instructors can set more complex criteria. They can add different options and criteria to rules. Students must meet all criteria in one of the rules to access them. Ten sections were applied during the first semester of the 2020-2021 academic year. The results proved their effectiveness.

TABLE 5: Comparison of metalearner.

Meta-learner	Accuracy (%)
MLP	97.3
RF	97.1
XGB	96.5
LGBM	96

TABLE 6: Performance metrics in % of the sgfp vs base classifiers.

Model	Accuracy	Precision	Recall	F1-score
LGBM	97.1	96.9	97.1	96.8
XGB	96.8	96.6	96.8	96.5
RF	96.3	96.8	96.3	96.4
SGFP	97.3	97.2	97.3	97.1

Boldface indicates the best result.

5. Results Analysis

The application of the SGFP approach on 10 sections during the first semester of the academic year 2020-2021 raised the following points (Table 8):

Before sitting on the midterm exam, the students were classified as follows: 28% in class A, 60% in class B, and 12% in class C. This means that a significant percentage of students were threatened by failure.

Three learning paths were customized to each students' class, and the results of the classification after passing the midterm exam were 36% in class A and 62.86% in class B, while class C contained only 1.14% (only two students).

5.1. SGFP Performances. The experiment of the SGFP algorithm on 10 sections of 176 students enrolled during the academic year 2020-2021, allowed the success of 174 students and the failure of only two; so, 98,86% of the students took the midterm exam. Moreover, the SGFP approach gave satisfactory results when tested with additional students enrolled in different colleges in Saudi universities (such as, college of nursing, college of business). Prior to the exam, instructors extract data from the Blackboard (Evaluation -course analytics-) via the 'Course at-a-Glance' option to generate input data for the SGFP approach, and all enrolled students are classified. Then, instructors confirm the testing and validation of the model. The particularity of this method lies in the fact that adequate learning of the SGFP model allows for an efficient classification of the student. The results show that the performance of SGFP is almost perfect. All the learning paths taken by the students were customized with the "Adaptive Release" tool of Blackboard and this according to the obtained grades in subject assessments. In fact, students were not able to move from one item to another only if they passed the test with a score of more than 70 percent. The Shareable Content Object Reference Model (SCORM) package is used to generate an individual learning path in the LMS (i.e., personalized contents navigating and sequence, assessments).

The results show SGFP approach performs better than one classifier learning model. F1-score, accuracy and precision values are higher compared to LGBM, XGB, and RF models.

TABLE 7: The validation dataset.

Course name	Number of students	Section no.	Faculty/department	Method of learning
Internet applications (T)	10	1	CCIS/ISD	Fully online
Internet applications (P)	10	1	CCIS/ISD	Fully online
Internet applications (T)	14	2	CCIS/ISD	Fully online
Internet applications (T)	14	2	CCIS/ISD	Fully online
Internet applications (P)	41	3	CCIS/ISD	Fully online
Internet applications (T)	41	3	CCIS/ISD	Fully online
Internet applications (P)	50	4	CCIS/ISD	Fully online
Internet applications (P)	38	5	CCIS/ISD	Fully online
Computer skills (T)	23	9	FCYC/CSD	Fully online
Computer skills (P)	23	9	FCYC/CSD	Fully online
Total of students	176			

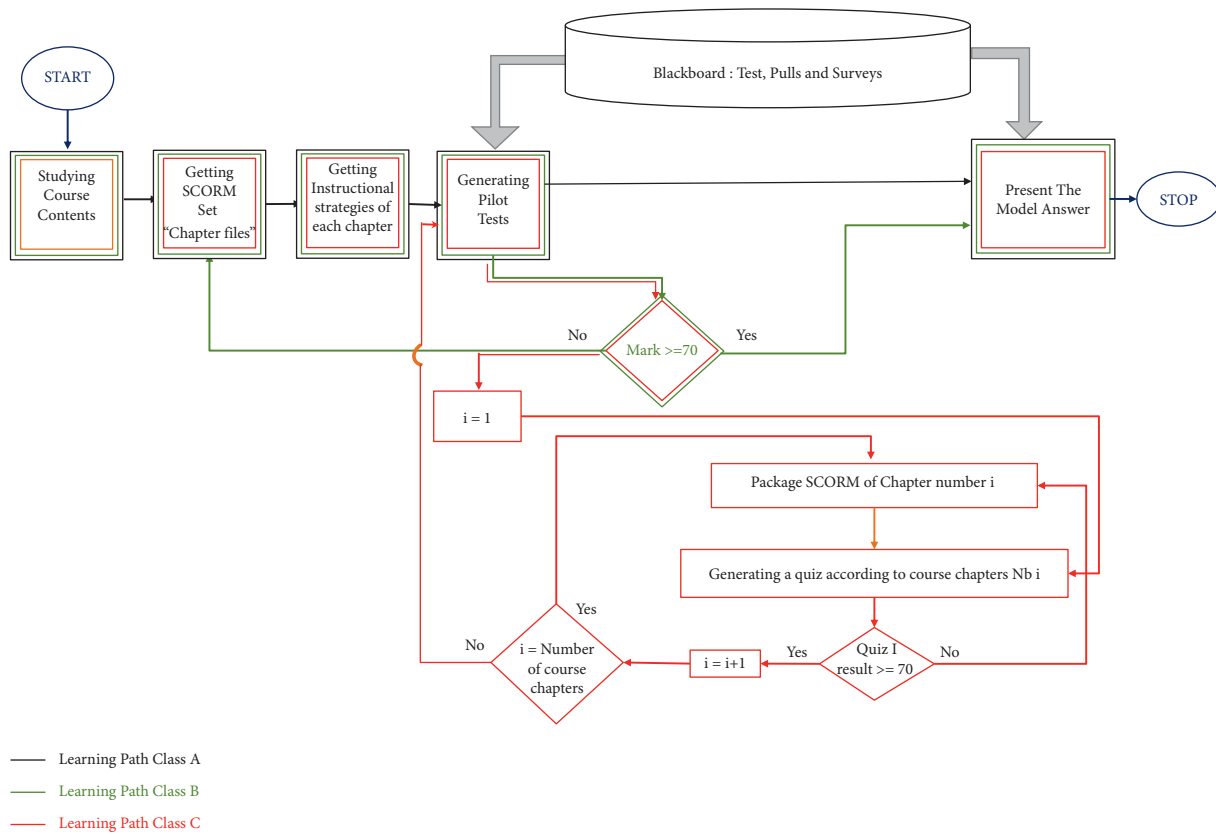


FIGURE 6: Learning paths.

TABLE 8: Distribution of classes' students before and after applying the learning paths.

Classes	Before using learning paths, % (number)	After using learning paths, % (number)
Class A	28% (49)	36% (63)
Class B	60% (106)	62.86%(111)
Class C	12%(21)	1.14%(2)

5.2. *Limitations.* Our promising SGFP results are data dependent. The dataset used for learning and testing has the distinction of being objective, clean, and reflective of student behaviour in the online environment. So, we suspect two major limitations related to:

- (i) Training/testing corpus size: using only one testing and two training semesters is not enough to maximize the SGFP performances. The recommendation here is to increase the number of the semesters but without overfitting the model.

- (ii) Dataset source: our approach uses LMS students' data. Even some students feel to be more confident to surveys data. The limitation here is that data related to students' behaviour can be interesting to consider. Thereby, an amalgamation of survey data and LMS data could increase precision and accuracy of SGFP approach.
- (iii) Exhaustiveness of attributes: here we use student university data encompassing only traditional online assessments (Access Operations, LMS Spent Time, Grade of Evaluation Center, etc.). We think this is not sufficient, and non-traditional attributes can be added to predict students' failure in online environment (i.e., students engagements variables, attendance, time point intervention, etc.). A comprehensive study including these variables will be considered in future work by adding more data from LMS platforms (i.e., students files logs, students' participations on discussion forums, blogs, wikis, etc.).

6. Conclusion

In this paper, a stacked generalisation-based algorithm (SGFP) is proposed to predict and avoid student failures using data from the analytical reports of learning management systems and grade containers of undergraduate students.

The robustness of this study is linked to three factors:

- (i) The data used to classify the students are overly significant
- (ii) The classification method is based on three prediction models rather than being content with a single model
- (iii) The automatic design of customized learning paths for each student depending on the class he belongs using LMS tools and this without intimidating the student

The first stage of this study was the classification of students according to their performance using data from A4L and FGC. The authors believe that to better predict student dropout, meaningful and objective data will be of great help. The SGFP algorithm is based on heterogeneous data that reflects students' levels of interactivity and their performance and teaching skills. These data are more relevant and beneficial to the e-learning process and assessment. Accuracy, objectivity, and precision were obtained by extracting updated data from actual student activities. Indeed, the results of algorithms/learning models in e-learning environments are more fruitful if the data used comes directly from LMS platforms.

The findings are three students' classes (Class A: above average, Class B: average, Class C: below average). The second stage is to customize learning paths (contents and assessments) according to students' classes by using the "Adaptive Release" tool capabilities of Blackboard. This tool is used to personalize learning routes and is based on a set of

rules regarding four different types of criteria: date, membership, grade, and review status. The customized content publication also aims to create courses that are more dynamic and appropriate for the needs of each learner.

The principal goal achieved is the improvement of students' success rate (98.86%). After using the SGFP model, students in class C will almost certainly fail (some students in class B are threatened by failure). The results show a significant improvement in students' success rates in the academic year 2020-2021. For class C students, the failure rate increases from 12% to 1.14%, for whom more evaluation steps and contents are customized. Following well-defined learning paths, only two out of 176 students failed. Therefore, the success rate was 98.86%.

The SGFP approach, which is built on a combination of an ensemble of failure prediction models, identified students' classes with a sensitivity average of 97.3% and a precision average of 97.2%. Compared to base classifiers, the results show that the mean and median accuracies of SGFP are higher. SGFP had the highest F1-score of 97.1%.

The robustness of this study to better predict failure is primarily linked to the good quality of the data, which are recent, significant, and objective. The learning phase from indicators extracted from e-learning environments is very effective because the data used comes directly from LMS platforms. This study comes to face the limitations of the previous works by presenting the idea of exploiting the analytical reports of the LMS which records all the activities of the students in detail and represents a clear mirror of each student. In fact, all LMS platforms provide an amalgamation of comprehensive and easy-to-use analytical reporting tools. These reports are interesting sources of knowledge to explore in adaptive learning. Concerning the classification robustness, the three models chosen had presented their satisfaction after several tests.

Among the recent publications, we notice that our study presented a clear improvement. Indeed, the precision of SGFP is 97.2%, while the best precision value obtained from [60] is 60.2%, achieved by the Hard-major classifier. The study in [62] reached 97% with the random forest classifier. A precision of 63.73% was achieved by LGBALD classifier in [63].

8F1-score achieved by SGFP is 97.1%, while the best value obtained for the published recent works in [26, 60, 62, 63] is 93.2% achieved by the J48 Decision Tree classifier.

Finally, the outcomes obtained in this study can be beneficial for higher institutions implementing a full distance-learning mode or in some pandemic health circumstances to improve student's performance and hiring rates in local and internal labor markets. This work is still open for improvements by (i) enhancing the stacked generalization training performances by combining rule-based classification and learning datasets, (ii) implementing adaptive learning according to learners' preferences and learning styles (i.e., visual learners, social learners, auditory learners, etc.), and (ii) adopting other classification algorithms and models and evaluating and comparing their results [17, 61, 64, 65].

6.1. *Ethic Statement.* This study did not involve human participants; there was no need for participant consent in this study, and no minors were included in this study. This study did not report medical records or archived samples.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Deanship of Scientific Research at Umm al-Qura University for supporting this work by grant code 18-EDU-1-01-0004.

References

- [1] R. S. Nickerson, "Technology in education in 2020: thinking about the not-distant future," in *Technology in Education*, pp. 19–28, Routledge, London, United Kingdom, 2013.
- [2] N. Selwyn, *Telling Tales on Technology: Qualitative Studies of Technology and Education*, Routledge, London, United Kingdom, 2020.
- [3] N. Hasan and Y. Bao, "Impact of "e-learning crack-up" perception on psychological distress among college students during covid-19 pandemic: a mediating role of "fear of academic year loss"," *Children and Youth Services Review*, vol. 118, Article ID 105355, 2020.
- [4] K. Naji and A. Ibriz, "Adaptive MOOC supports the elicitation of learners' preferences," in *International Conference on Advanced Intelligent Systems for Sustainable Development*, pp. 68–73, Springer, Berlin, Germany, 2019.
- [5] H. Wang, "College physical education and training in big data: a big data mining and analysis system," *Journal of Healthcare Engineering*, vol. 2021, Article ID 3585630, 8 pages, 2021.
- [6] M. Ragab, A. M. Abdel Aal, A. O. Jifri, and N. F. Omran, "Enhancement of predicting students performance model using ensemble approaches and educational data mining techniques," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6241676, 2021.
- [7] H. Khosravi, S. Sadiq, and D. Gasevic, "Development and adoption of an adaptive learning system: reflections and lessons learned," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, pp. 58–64, Portland, OR, USA, March 2020.
- [8] G. C. Crisan, "From digital learning resources to adaptive learning objects: an overview," *Modelling and Development of Intelligent Systems*, in *Proceedings of the 6th International Conference, MDIS 2019*, p. 18, Sibiu, Romania, October 2019.
- [9] L. Liu, "Design and implementation of English listening teaching based on a wireless communication microprocessor and virtual environment," *Journal of Sensors*, vol. 2021, Article ID 2887302, 2021.
- [10] Y. Jin, Y. Yang, B. Yang, and Y. Zhang, "An adaptive BP neural network model for teaching quality evaluation in colleges and universities," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 4936873, 2021.
- [11] F. E. Louhab, A. Bahnasse, F. Bensalah, A. Khiat, Y. Khiat, and M. Talea, "Novel approach for adaptive flipped classroom based on learning management system," *Education and Information Technologies*, vol. 25, no. 2, pp. 755–773, 2020.
- [12] S. Shekapure and D. D. Patil, "Learning preferences analysis by case-based reasoning," in *Computing in Engineering and Technology*, pp. 145–153, Springer, Berlin, Germany, 2020.
- [13] M. Cerna, "Modified recommender system model for the utilized elearning platform," *Journal of Computers in Education*, vol. 7, no. 1, pp. 105–129, 2020.
- [14] L. Liu, H. Jiang, P. He et al., "On the variance of the adaptive learning rate and beyond," arXiv preprint arXiv:1908.03265, 2019.
- [15] E. M. Smaili, S. Sraidi, S. Azzouzi, and M. E. H. Charaf, "Towards sustainable e-learning systems using an adaptive learning approach," in *Emerging Trends in ICT for Sustainable Development*, pp. 365–372, Springer, Berlin, Germany, 2021.
- [16] C. Wang, "Analysis of students' behavior in english online education based on data mining," *Mobile Information Systems*, vol. 2021, Article ID 1856690, 10 pages, 2021.
- [17] L. Q. Chen, M. T. Wu, L. F. Pan, and R. B. Zheng, "Grade prediction in blended learning using multisource data," *Scientific Programming*, vol. 2021, Article ID 4513610, 2021.
- [18] J. Liu, C. Yin, Y. Li, H. Sun, and H. Zhou, "Deep learning and collaborative filtering-based methods for students' performance prediction and course recommendation," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2157343, 2021.
- [19] X. Shen and C. Yuan, "A college student behavior analysis and management method based on machine learning technology," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 3126347, 2021.
- [20] M. Adnan, A. Habib, J. Ashraf et al., "Predicting at-risk students at different percentages of course length for early intervention using machine learning models," *IEEE Access*, vol. 9, pp. 7519–7539, 2021.
- [21] A. F. Núñez-Naranjo, M. Ayala-Chauvin, and G. Ribas-Sanmartí, "Prediction of university dropout using machine learning," in *International Conference on Information Technology & Systems*, pp. 396–406, Springer, Berlin, Germany, 2021.
- [22] O. Sukhbaatar, T. Usagawa, and L. Choimaa, "An artificial neural network based early prediction of failure-prone students in blended learning course," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 14, no. 19, pp. 77–92, 2019.
- [23] W. F. Wan Yaacob, N. Mohd Sobri, S. A. M. Nasir, W. F. Wan Yaacob, N. D. Norshahidi, and W. Z. Wan Husin, "Predicting student drop-out in higher institution using data mining techniques," *Journal of Physics: conference Series*, vol. 1496, no. 1, Article ID 012005, 2020.
- [24] Y. Zhang, R. An, J. Cui, and X. Shang, "Undergraduate grade prediction in Chinese higher education using convolutional neural networks," in *Proceedings of the LAK21: 11th International Learning Analytics and Knowledge Conference*, pp. 462–468, New York, NY, USA, April 2021.
- [25] M. Ababneh, A. Aljarrah, D. Karagozlu, and F. Ozdamli, "Guiding the students in high school by using machine learning," *TEM Journal*, vol. 10, no. 1, pp. 384–391, 2021.
- [26] S. B. Keser and S. Aghalarova, "Hela: a novel hybrid ensemble learning algorithm for predicting academic performance of students," *Education and Information Technologies*, pp. 1–32, 2021.

- [27] S. Zian, S. A. Kareem, and K. D. Varathan, "An empirical evaluation of stacked ensembles with different meta-learners in imbalanced Classification," *IEEE Access*, vol. 9, 2021.
- [28] A. Siri, "Predicting students' dropout at university using artificial neural networks," *Italian Journal of Sociology of Education*, vol. 7, no. 2, 2015.
- [29] A. Graves, S. Fernandez, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks," in *Proceedings of the 23rd International Conference on Machine Learning*, pp. 369–376, Pittsburgh, Pennsylvania, June 2006.
- [30] G. Hepner, T. Logan, N. Ritter, and N. Bryant, "Artificial neural network classification using a minimal training set-comparison to conventional supervised classification," *Photogrammetric Engineering & Remote Sensing*, vol. 56, no. 4, pp. 469–473, 1990.
- [31] R. Cl erot and F. Clerot, "A methodology to explain neural network classification," *Neural Networks*, vol. 15, no. 2, pp. 237–246, 2002.
- [32] B. Asadi and H. Jiang, "On approximation capabilities of relu activation and softmax output layer in neural networks," arXiv preprint arXiv:2002.04060, 2020.
- [33] H. Waheed, M. Anas, S.-U. Hassan et al., "Balancing sequential data to predict students at-risk using adversarial networks," *Computers & Electrical Engineering*, vol. 93, Article ID 107274, 2021.
- [34] R. Alamri and B. Alharbi, "Explainable student performance prediction models: a systematic review," *IEEE Access*, vol. 9, Article ID 3061368, 2021.
- [35] N. Chanamarn, K. Tamee, and P. Sittidech, "Stacking technique for academic achievement prediction," in *Proceedings of the International Workshop on Smart Info-Media Systems in Asia (SISA 2016)*, pp. 14–17, Ayutthaya, Thailand, September 2016.
- [36] B. S. Kalyani, D. Harisha, V. RamyaKrishna, and S. Manne, "Evaluation of students performance using neural networks," in *International Conference on Intelligent Computing, Information and Control Systems*, pp. 499–505, Springer, Berlin, Germany, 2019.
- [37] F. Okubo, T. Yamashita, A. Shimada, and H. Ogata, "A neural network approach for students' performance prediction," in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, pp. 598–599, New York, NY, USA, March 2017.
- [38] B. Anthony, A. Kamaludin, A. Romli et al., "Blended learning adoption and implementation in higher education: a theoretical and systematic review," *Technology, Knowledge and Learning*, pp. 1–48, 2020.
- [39] J. Willison, "Blended learning needs blended evaluation," in *Critical Perspectives on Teaching, Learning and Leadership*, pp. 87–106, Springer, Berlin, Germany, 2020.
- [40] W. S. Chen and A. Y. Tat Yao, "An empirical evaluation of critical factors influencing learner satisfaction in blended learning: a pilot study," *Universal Journal of Educational Research*, vol. 4, no. 7, pp. 1667–1671, 2016.
- [41] S. Keskin, "Factors affecting students' preferences for online and blended learning: motivational vs. cognitive," *European Journal of Open, Distance and E-Learning*, vol. 22, no. 2, pp. 72–86, 2019.
- [42] R. Shenbagaraj and S. Iyer, "A smart education solution for adaptive learning with innovative applications and interactivities," in *Emerging Technologies in Data Mining and Information Security*, pp. 691–697, Springer, Berlin, Germany, 2021.
- [43] L. L. W. Cheung and A. C. N. Kan, "Evaluation of factors related to student performance in a distance-learning business communication course," *Journal of Education for Business*, vol. 77, no. 5, pp. 257–263, 2002.
- [44] Y. Li, Z. Shao, X. Wang, X. Zhao, and Y. Guo, "A concept map-based learning paths automatic generation algorithm for adaptive learning systems," *IEEE Access*, vol. 7, pp. 245–255, 2018.
- [45] G. Jagadamba, R. Sheeba, K. Brinda, K. Rohini, and S. Pratik, "Adaptive e-learning authentication and monitoring," in *Proceedings of the 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 277–283, IEEE, Bangalore, India, March 2020.
- [46] E. Osmanbegovic and M. Suljic, "Data mining approach for predicting student performance," *Economic Review: Journal of Economics and Business*, vol. 10, no. 1, pp. 3–12, 2012.
- [47] M. S. Kumar and G. P. Babu, "Comparative study of various supervised machine learning algorithms for an early effective prediction of the employability of students," *Journal of Engineering Sciences*, vol. 10, no. 10.
- [48] A. M. Shahiri, W. Husain, and N. Rashid, "A review on predicting student's performance using data mining techniques," *Procedia Computer Science*, vol. 72, pp. 414–422, 2015.
- [49] R. Hasan, S. Palaniappan, S. Mahmood, A. Abbas, K. U. Sarker, and M. U. Sattar, "Predicting student performance in higher educational institutions using video learning analytics and data mining techniques," *Applied Sciences*, vol. 10, no. 11, p. 3894, 2020.
- [50] S. Kotsiantis, C. Pierrakeas, and P. Pintelas, "Predicting students' performance in distance learning using machine learning techniques," *Applied Artificial Intelligence*, vol. 18, no. 5, pp. 411–426, 2004.
- [51] L. Smirani, J. Boulahia, and R. Bouallegue, "A semi blind channel estimation method based on hybrid neural networks for uplink LTE-A," *International Journal of Wireless & Mobile Networks*, vol. 8, 2017.
- [52] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 23, Article ID e5909, 2021.
- [53] S. Lobov, V. Mironov, I. Kastalskiy, and V. Kazantsev, "A spiking neural network in semg feature extraction," *Sensors*, vol. 15, no. 11, pp. 27 894–27 904, 2015.
- [54] K. Ghefiri, S. Bouall egue, I. Garrido, A. Garrido, and J. Hagg ege, "Multi-layer artificial neural networks based mppt-pitch angle control of a tidal stream generator," *Sensors*, vol. 18, no. 5, p. 1317, 2018.
- [55] P. M. Arsad and N. Buniyamin, "A neural network students' performance prediction model (nnsppm)," in *Proceedings of the 2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*, pp. 1–5, IEEE, Kuala Lumpur, Malaysia, November 2013.
- [56] A. M. Adewale, A. O. Bamidele, and U. O. Lateef, "Predictive modelling and analysis of academic performance of secondary school students: artificial neural network approach," *International Journal of Science and Technology Education Research*, vol. 9, no. 1, pp. 1–8, 2018.
- [57] A. R. Iyanda, O. D. Ninan, A. O. Ajayi, and O. G. Anyabolu, "Predicting student academic performance in computer science courses: a comparison of neural network models," *International Journal of Modern Education and Computer Science*, vol. 10, no. 6, 2018.

- [58] R. Ahuja and S. Sharma, "Stacking and voting ensemble methods fusion to evaluate instructor performance in higher education," *International Journal on Information Technology*, vol. 13, pp. 1–11, 2021.
- [59] A. Alizamar, Y. Syahputra, A. Afdal, Z. Ardi, and L. Trizeta, "Differences in aggressive behavior of male and female students using rasch stacking," *International Journal of Research in Counseling and Education*, vol. 3, no. 1, pp. 22–32, 2018.
- [60] C. S. Rao and A. S. Arunachalam, "Ensemble based learning style identification using VARK," *NVEO-Natural Volatiles & Essential OILS Journal| NVEO*, pp. 4550–4559, 2021.
- [61] A. Onan and M. A. Tocoglu, "A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification," *IEEE Access*, vol. 9, pp. 7701–7722, 2021.
- [62] D. Aggarwal, S. Mittal, and V. Bali, "Significance of non-academic parameters for predicting student performance using ensemble learning techniques," *International Journal of System Dynamics Applications*, vol. 10, no. 3, pp. 38–49, 2021.
- [63] L. Zhang, S. Kai, H. Keyu, and Z. Ruiqiu, "An approximation of label distribution-based ensemble learning method for online educational prediction," *International Journal of Computers, Communications & Control*, vol. 16, no. 3, 2021.
- [64] A. Onan, "Sentiment analysis on massive open online course evaluations: a text mining and deep learning approach," *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 572–589, 2021.
- [65] A. Onan, "Ensemble of classifiers and term weighting schemes for sentiment analysis in Turkish," *Scientific Research Communications*, vol. 1, no. 1, 2021.

Research Article

Research on Teaching Reform of College Student Training Mode Based on Engineering Project Economic Evaluation of Driving Behavior with Internet of Vehicles Data

Shuilong He ^{1,2}, Zhihong Zou ¹, Hui Li ¹, Jucai Deng ², Enyong Xu ²,
Rongjiang Tang ¹ and Yanyan Zhou ¹

¹School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin, 541004, China

²Dongfeng Liuzhou Motor Co., Ltd., Liuzhou 545005, China

Correspondence should be addressed to Rongjiang Tang; 251444972@qq.com and Yanyan Zhou; 190835141@qq.com

Received 1 November 2021; Revised 3 January 2022; Accepted 12 March 2022; Published 12 April 2022

Academic Editor: Chenxi Huang

Copyright © 2022 Shuilong He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Practice is one of the essential teaching links in application-oriented professional teaching in engineering colleges. Reasonable design of practical teaching mode has an important influence on the development of scientific research activities and training of applied talents. In order to better experience the learning combined college student teaching mode, the analysis and mining based on the Internet of Vehicles data is taken as the research scene. Theoretical research and engineering verification of drivers' driving behavior economy are carried out by using the learning method of theoretical research under the guidance of teachers and engineering practice under the guidance of enterprises. An economic evaluation model and energy saving potential calculation method based on fuzzy analytic hierarchy process are established, and the model is verified and improved in engineering practice. Among them, the analysis of personalized characteristics of driver behavior indicators shows that there are obvious differences in individual preference characteristics, generally manifested as fast acceleration and deceleration, low speed driving, gear mismatch. In addition, some drivers' bad driving behavior has an energy saving potential of up to 4.98%. The results show that the combination of school theory research and enterprise engineering practice has positive effects on the development of students' scientific research and the cultivation of applied talents that contribute to the development of enterprises.

1. Introduction

China Association for Professional Certification of Engineering Education defines an authoritative index of engineering education in colleges and universities, namely the engineering education certification standard. The standard requires students majoring in engineering to complete relevant courses, engineering practice and graduation projects to achieve graduation requirements. This for the school and our students both put forward the theoretical requirements also put forward the practical requirements. Therefore, we should give consideration to both theoretical research and engineering practice in our study and life. In order to better participate in the new model of joint teaching of theoretical research and engineering practice, relevant

research is carried out. The research scenario is based on the analysis and mining of Internet of vehicles data. Theoretical research and engineering verification are carried out on the economic evaluation of drivers' driving behavior by using the learning method of theoretical research under the guidance of teachers and engineering practice under the guidance of enterprises. Firstly, literature research was carried out. Secondly, theoretical research was conducted under the guidance of teachers. Finally, engineering practice and theoretical improvement were carried out with the help of enterprises. The specific research process is described as follows.

In recent years, with the development of transportation and logistics industry, the amount of road cargo transport is increasing, and commercial vehicles are playing an

important role in the transport of goods. By the end of 2019, China will have 10,878,200 trucks. However, commercial vehicles, which account for 10.9% of the total automobile volume, consume much more fuel than passenger vehicles and emit 70% nitrogen oxide (NO_x) and more than 90% particulate matter (PM). It not only brings huge fuel consumption costs [1], but also exerts a serious impact on the atmospheric environment. Therefore, improving the fuel efficiency of vehicles and reducing fuel consumption of vehicles are urgent problems to be solved in the current transportation field. It has an important influence on the vehicle to meet the requirement of “energy saving and emission reduction”.

Many scholars have studied this problem. Among them, improving engine “thermal management” technology [2], improving transmission efficiency [3] and vehicle lightweight research [4] are the main approaches to improve fuel efficiency and reduce fuel consumption in traditional research. With the rise of ecological driving concept, optimizing driving behavior has become another important direction to improve fuel efficiency and reduce fuel consumption. Literature [5] found that in aggressive driving operations, fuel consumption increased by 12%–40%, emissions increased by 20%–50%. Literature [6] indicates that fuel consumption can be significantly reduced by 5% ~ 25% by using energy-saving driving operations. The results show that different driving styles have significant effects on fuel consumption and emissions. Krishnamoorthy and Gopalakrishna [7] studied the method of evaluating the driving ability of truck drivers. Using the data provided by the fleet management system, the influences of driving behavior factors on fuel consumption, such as acceleration at high speed, long idle speed, overspeed, gear speed mismatch and engine noneconomic speed, are analyzed. Frank et al. [8] built an Android application to assist driving behavior. By collecting data related to the car’s CAN bus, the driver CAN obtain a representative ecological score per second. The system can introduce the basic concepts and suggestions of green driving to the driver during driving. Fuel consumption tests conducted by seven volunteers showed that the Android app significantly reduced overall energy consumption. Rolim [9] studied the influence of real-time feedback on ecological driving behavior and the variables affecting fuel consumption. Data analysis by The Lisbon bus operator shows that in the absence of real-time feedback on driving behavior, the number of incidents of bad driving behavior has increased significantly. At the same time, trends in fuel consumption were similar to bad driving behavior. Ferreira et al. [10] studied the impact of driving style on fuel consumption by using the speed, acceleration, engine speed and other parameters collected by car CAN bus and GPS device. To assess the level of driving behavior, divide driving behavior into 5 categories. Hsu [11] established an ecological driving behavior analysis model for driving decision-making through data mining technology in order to improve driving efficiency. Aiming at the influence of driver’s personal behavior and vehicle type on driving efficiency, a new comprehensive driving efficiency index was proposed to evaluate driving behavior. Hoang [12] pointed out that improving the

service level is an inevitable requirement for the sustainable development of urban public transport, and proposed a new model to evaluate the driving behavior of bus drivers through traffic planning. In addition, there are many studies in this field [13–18], which prove the feasibility of eco-driving technology to improve the fuel efficiency of existing vehicles. It provides a theoretical basis for further optimization of driving behavior.

In these studies, although the influence of driving behavior indicators on fuel consumption and its relative influence degree are discussed, there is little research on the preference characteristics and energy saving potential of drivers in many bad driving behaviors. However, the analysis of drivers’ preference characteristics in many bad driving behaviors and the accurate description of their fuel consumption impact have important influence on the formulation of efficient and personalized energy-saving optimization strategies. To sum up, this paper carries out relevant work in view of the deficiencies of existing studies. First of all, based on the data of commercial vehicles connected to the Internet, relevant analysis method is adopted to select the economic indicators of driving behavior that have an impact on fuel consumption, and further carry out quantitative evaluation research on driving behavior economy. Secondly, the personalized characteristics of drivers in many bad driving behaviors are analyzed based on the economic evaluation results. Finally, the driving behavior corresponds to the energy saving potential calculation.

The thesis is briefly summarized as follows. In section 2, the theories used in driving behavior evaluation are introduced. Section 3 discusses the process of driving behavior economy evaluation modeling. Section 4 Test the system through real car data and analyze the test results. Section 5 summarizes the work of the full text, and further analyzes the existing shortcomings and prospects.

2. Theory of Economic Evaluation of Driving Behavior

2.1. Selection of Driving Behavior Indicators. In order to reasonably construct the evaluation system of driving behavior economy, the selected evaluation indexes should be targeted and comprehensive. Studies have shown that there are many driving behavior indicators that have an impact on fuel consumption, and the impact degree of relative fuel consumption varies among different indicators. Literature [19] screened driving behavior evaluation indicators by starting from the duration of driving events and combining with data variables representing driving behavior. Literature [10, 20, 21] explored the influence degree of different driving behavior indicators on fuel consumption based on Naive Bayes method, correlation analysis and other methods, and selected driving behavior indicators based on the relative influence degree. The methods and results of index selection of all literature were summarized, and the indexes were preliminarily selected. In order to further analyze the impact of selected indicators on the target fuel consumption of 100 km, Spearman correlation coefficient method [21] was used to calculate the correlation between preliminarily

selected driving behavior indicators and fuel consumption of 100 km, as shown in equation (1). The rationality of index selection is tested by correlation coefficient. ρ is Spearman correlation coefficient, X_i and Y_i are the index data involved in the calculation. N is the dimension of the data.

$$\rho = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2 \sum_{i=1}^N (Y_i - \bar{Y})^2}} \quad (1)$$

The calculation results of correlation coefficients are shown in Table 1, where the mean speed, mean speed and fuel consumption of the index engine are less than 0.1. It shows that the correlation between these indexes and 100 km fuel consumption is low and can be ignored. Therefore, after removing these irrelevant indicators, 11 indicators with relatively large correlation are finally selected for economic evaluation. To simplify index variable names, abbreviations are used instead of the original names. The details are as follows: ratio of driving idle speed (DS), the proportion of time spent driving at low velocity (LS), speed standard deviation (VSTD), acceleration time ratio (HA), acceleration mean (AMEAN), engine speed standard deviation (SSTD), deceleration time ratio (LA), acceleration rate of change (AC), large throttle operating time ratio (DY), overspeed time percentage (CS), high speed/low speed ratio (LDHV), average engine speed (SMEAN), and average engine speed (VMEAN).

2.2. Establish Evaluation System

2.2.1. Construction of Driving Behavior Economy Evaluation System. Based on the above selected economic evaluation indexes of driving behavior, analytic hierarchy process (AHP) is applied to subdivide the index levels step by step to establish a ladder system of driving behavior evaluation [22]. The target layer is the driving behavior economy score, and the criterion layer is the index of each level. The first-level index U_i has economic evaluation variables such as engine speed, gearbox output shaft speed, speed and acceleration. The second-level index U_{ij} is the statistic corresponding to the economic evaluation variable. The solution layer is each mileage segment, as shown in Figure 1.

2.2.2. Establish the Relationship between Driving Behavior Indicators and Loss Score. Different drivers operate their vehicles differently, resulting in different fuel consumption and economy scores. Relative to the correct operating mode, the actual operation of each driving economy evaluation index X_i will bring driving efficiency rating score loss value ΔF_i and fuel consumption loss ΔQ_i . This process can be described by equations (2), (3).

Here f_i and q_i are the loss description functions corresponding to the i -level index.

$$\Delta F_i = f_i(X_i), \quad (2)$$

$$\Delta Q_i = q_i(X_i). \quad (3)$$

In addition, it is characterized by uncertainty and fuzziness in the process of judging the comparison between actual driving behavior and correct driving behavior. Therefore, it provides an idea for establishing the actual driving behavior economy evaluation model and energy saving potential analysis.

2.3. Fuzzy Comprehensive Evaluation Algorithm. Based on the process analysis of the influence of driving behavior on fuel consumption and the process analysis of driving behavior evaluation, the paper uses fuzzy comprehensive evaluation algorithm to evaluate driving behavior economy [23]. The theory and process of fuzzy evaluation algorithm are introduced below.

2.3.1. Determine the Set of Evaluation Factors. The evaluation factor set is determined according to the evaluation objective. The factors of establishing the driving behavior economy evaluation system based on AHP are as follows. U is the set containing all evaluation factors, and the set element is the driving behavior evaluation factor.

$$U = \{U_1, U_2, \dots, U_m\}. \quad (4)$$

2.3.2. Establish the Comment Set. The grade is divided, and the comment sets corresponding to different grades are determined. V is the set of all comment levels. The set element is the rating level. Fen is the rating set of the corresponding comment and the element is the subset of the rating.

$$\begin{aligned} V &= \{V_1, V_2, \dots, V_n\}, \\ fen &= \{fen_1, fen_2, \dots, fen_n\}. \end{aligned} \quad (5)$$

2.3.3. Determine the Weight Vector of Evaluation Index. Determine the relative influence degree of evaluation factors on the evaluation target, namely the weight W . Evaluation index weight vector W_{ij} corresponds to the weight of the j second-level evaluation index under the i first-level evaluation index.

$$W = \{w_{11}, w_{12}, w_{13}, w_{21}, w_{22}, w_{23}, w_{24}, w_{31}, w_{41}, w_{42}, w_{43}\}. \quad (6)$$

2.3.4. Construct Fuzzy Relation Matrix. The fuzzy membership function is selected based on the data distribution characteristics of the index, and the selected membership function is used to construct the fuzzy relation matrix. R_{ik} is the membership value corresponding to the i driving behavior economy index at different evaluation levels. R_{ik} is combined into a single-index membership vector R_i , and all the single-index membership vectors are combined into a comprehensive evaluation membership matrix R .

TABLE 1: Driving behavior index and fuel consumption correlation coefficient.

Indicator name abbreviation	Spearman correlation coefficient (ρ)	Correlation ranking
DS	0.303**	1
LS	0.306**	2
VSTD	0.255**	3
HA	0.226**	4
AMEAN	0.205**	5
SSTD	0.195**	6
LA	0.193**	7
AC	0.171**	8
DY	0.155**	9
CS	0.151**	10
LDHV	0.104**	11
SMEAN	0.052*	12
VMEAN	-0.049*	13

**indicates significant correlation at 0.01 level.

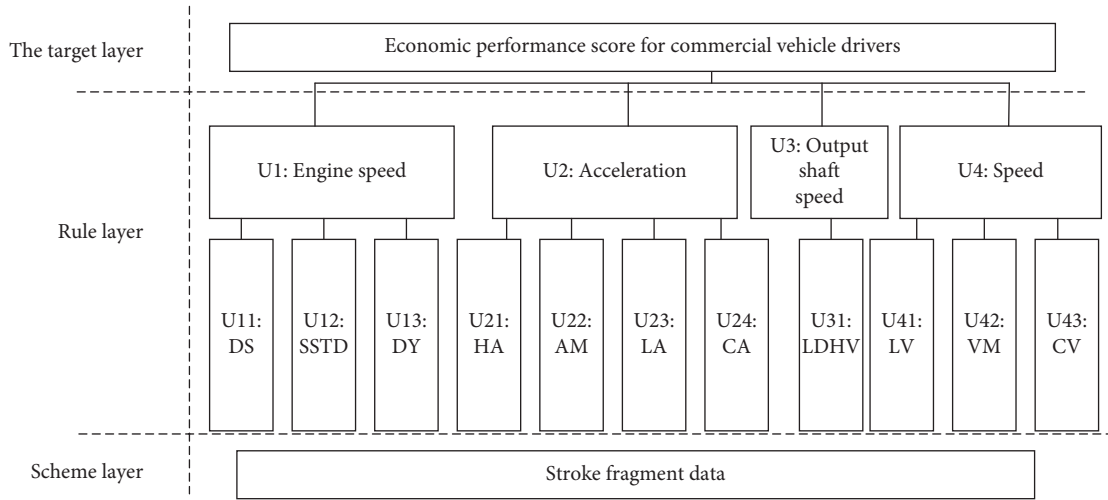


FIGURE 1: Driving behavior economy system evaluation system.

$$R_i = [R_{i1}, R_{i2}, \dots, R_{ik}, \dots, R_{in}],$$

$$R = \begin{bmatrix} R_1 \\ R_2 \\ \cdot \\ \cdot \\ \cdot \\ R_m \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & \cdot & \cdot & R_{1n} \\ R_{21} & R_{22} & \cdot & \cdot & R_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ R_{m1} & R_{m2} & \cdot & \cdot & R_{mn} \end{bmatrix}. \quad (7)$$

2.3.5. *Fuzzy Calculation.* Through (8), the weight vector obtained and the membership matrix of multi-factor evaluation are fuzzy calculated. S is the obtained fuzzy evaluation result matrix, W is the determined index weight vector, R is the multi-factor evaluation membership matrix determined

by the formula, and the operation symbol is the weighted average fuzzy operator [23]. Formula (9) is used to normalize the fuzzy evaluation result matrix and sum it with the corresponding score segment to get the economic score F .

$$S = W \cdot \oplus R, \quad (8)$$

$$F = \sum_{i=1}^n \frac{S_i}{\sum_{i=1}^n S_i} \times f_i. \quad (9)$$

3. Modeling of Driving Behavior Economy Evaluation System Based on FAHP Algorithm

3.1. *Determine the Set of Evaluation Factors.* According to the hierarchy of driving behavior economy evaluation system in Figure 1, the set of factors is determined as follows, in which

U_{ij} corresponds to the j second-level evaluation index under the first-level evaluation index.

$$U = \{U1, U2, U3, U4\} = \left\{ \begin{array}{c} \{U_{11}, U_{12}, U_{13}\} \\ \{U_{21}, U_{22}, U_{23}, U_{24}\} \\ \{U_{31}\} \\ \{U_{41}, U_{42}, U_{43}, U_{44}\} \end{array} \right\}. \quad (10)$$

3.2. Comment Sets and Corresponding Scores. For each driving behavior indicator, the rating is divided and the corresponding single driving behavior loss assessment set and score set are given. The driving behavior economy score is obtained by comprehensively considering all loss scores, and the corresponding comment set and score set are shown in Tables 2, 3.

3.3. Determine the Weight of Evaluation Index. Different driving behavior indicators have different influences on driving behavior economy score and have different weights. By combining expert experience and 1–9 scale method [22], the judgment matrix showing the relative importance of economic evaluation indexes of driving behavior was given.

$$CI = \frac{\lambda_{\max} - b}{b - 1}, \quad (11)$$

$$CR = \frac{CI}{RI}. \quad (12)$$

In order to further calculate the weight vector, the judgment matrix should meet the requirement of consistency test. The judgment matrix is tested through equations (9) and (10), where CI is the consistency index, RI is the average consistency index, RI values correspond to different judgment matrix orders are shown in Table 4. Lambda Max is the maximum eigenvalue of the judgment matrix, b is the order of the judgment matrix, and CR is the consistency ratio.

The CR value of the judgment matrix is calculated successively, when the consistency ratio $CR < 0.1$, it indicates that the judgment matrix meets the consistency test requirements. If not, the judgment matrix is modified until it meets the requirements. The geometric average method is used to calculate the subjective weight W^a for the judgment matrix meeting the requirements as shown in equation (13), in which A_i is the judgment matrix and the element contained in the judgment matrix is a_{ij} .

$$W_i = \frac{(\prod_{j=1}^b a_{ij})^{1/b}}{\sum_{i=1}^m (\sum_{j=1}^b a_{ij})^{1/b}}, \quad (13)$$

$$W^a = \{W_1, W_2, \dots, W_m\}. \quad (14)$$

Considering the limitation of subjective weighting in analytic hierarchy process (AHP), the objective weighting

TABLE 2: Behavioral loss score set.

Collection of behavior loss comments	Behavior loss score set
Normal loss (I)	(-5,0)
Slight loss (II)	(-40, -5)
General loss (III)	(-60, -40)
A lot of loss (IV)	(-80, -60)
Severe losses (V)	(-100, -80)

TABLE 3: Behavioral evaluation score set.

Behavioral evaluation comment set	Behavior evaluation score set
Poor	(0,20)
Is very poor	(20,40)
Generally poor	(40,60)
Slightly less	(60,95)
Good	(95,100)

TABLE 4: Mean consistency test RI values.

Judge the order of the matrix(b)	1	2	3	4	5	6	7
RI values	0.00	0.00	0.58	0.9	1.12	1.24	1.32

method combined with correlation coefficient is used to improve subjective weighting. The correlation coefficient calculated by formula (1) reflects the objective characteristics of index data and fuel consumption, which can be used as the objective weight basis. The objective weight vector W_b is obtained after normalization. To sum up, the final weight value W is obtained by combining subjective weight and objective weight, as shown in Figure 2. $W_{U_{ij}}$ is the weight value corresponding to the index U_{ij} .

3.4. Construct Fuzzy Relation Matrix. To construct the fuzzy relation matrix, it is necessary to select the membership function with the same rule from the existing membership distribution function according to the characteristics of variables to characterize the fuzziness [24]. Data based on index variables have two characteristics: continuity and different influence degree of variable indexes on economic performance score. Different membership functions and membership parameters are used respectively. For the indexes that have great influence on economic evaluation, the fuzzy attribute changes sharply near the cut-off point of membership interval, so it is suitable to choose K parabolic membership function to describe this characteristic. Compare the relative weight vectors of the above driving behavior economy indicators, among which the indicators with larger weight are: idle time ratio, large throttle ratio, standard deviation of engine speed, and low speed driving ratio. Therefore, k-order parabolic type is selected as the membership function to describe the fuzziness. The fuzzy attributes of other variables change gently near the cut-off point, so it is appropriate to select the membership function as ridge function.

The expression of the membership function of the parabolic type of K is shown in equations (15)–(17), which

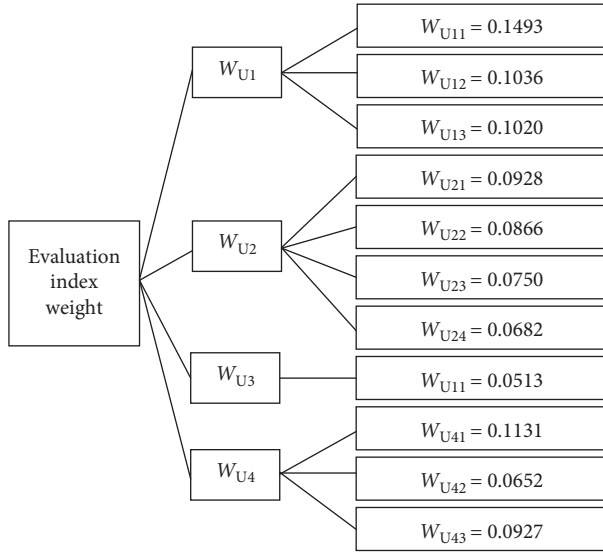


FIGURE 2: Index weight vector allocation.

are, in turn, relatively small, medium and large. Y is the corresponding membership value of each grade. X represents the index data, and X_i represents the threshold value of the index.

$$Y = \begin{cases} 1, & t < x_1, \\ \left(\frac{(x_2 - t)^2}{x_2 - x_1} \right)^2, & x_1 \leq t \leq x_2, \\ 0, & t \geq x_2, \end{cases} \quad (15)$$

$$Y = \begin{cases} 0, & t \leq x_1, \\ \left(\frac{(t - x_1)^2}{x_2 - x_1} \right)^2, & x_1 < t < x_2, \\ 1, & x_2 \leq t \leq x_3, \\ \left(\frac{(x_4 - t)^2}{x_4 - x_3} \right)^2, & x_3 < t < x_4, \\ 0, & t \geq x_4, \end{cases} \quad (16)$$

$$Y = \begin{cases} 0, & t \leq x_3, \\ \left(\frac{(t - x_3)^2}{x_4 - x_3} \right)^2, & x_3 < t < x_4, \\ 1, & t \geq x_4. \end{cases} \quad (17)$$

The ridge function expressions are shown in equations (18)–(20), which in turn are relatively small, intermediate and large.

$$Y = \begin{cases} 1, & t \leq x_1, \\ \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{x_2 - x_1} \left(t - \frac{x_1 + x_2}{2} \right), & x_1 < t < x_2, \\ 0, & t \geq x_2, \end{cases} \quad (18)$$

$$Y = \begin{cases} 0, & t \leq x_1, \\ \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{x_2 - x_1} \left(t - \frac{x_1 + x_2}{2} \right), & x_1 < t < x_2, \\ 1, & x_2 \leq t \leq x_3, \\ \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{x_4 - x_3} \left(t - \frac{x_4 + x_3}{2} \right), & x_3 < t < x_3, \\ 0, & t \geq x_4, \end{cases} \quad (19)$$

$$Y = \begin{cases} 1, & t \leq x_3, \\ \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{x_4 - x_3} \left(t - \frac{x_3 + x_4}{2} \right), & x_3 < t < x_4, \\ 0, & t \geq x_4. \end{cases} \quad (20)$$

In order to determine the parameter values of the membership function mentioned above, a statistical analysis of the characteristics was carried out based on a large number of real vehicle transportation data, and the parameters of the membership function obtained are shown in Table 5.

4. Engineering Practice Verification and Result Analysis of Evaluation Model

4.1. Driving Behavior Economy Evaluation Data Processing. The data in this paper comes from a T-box (vehicle-mounted intelligent terminal) that is the data acquisition terminal of commercial vehicles, and the interface is used to read the longitude and latitude, speed and instantaneous fuel consumption of the vehicle, etc. The data that can be uploaded by the acquisition terminal include: data acquisition time, vehicle identification code, ECU speed (km/h), instantaneous fuel consumption (L/100 KM), GPS longitude and latitude, cumulative mileage (KM), engine speed (R/min), engine torque (N·M), etc. A total of 36 vehicles of a freight company were collected within a week of the actual road traffic data.

The evaluation model is tested based on the transportation data of Internet of vehicles. To ensure the uniformity of index calculation, set mileage as the control variable. By dividing each driver's transport data into a uniform 20 KM Micro travel segment: a total of 10037 data segments were obtained by 36 drivers within one week, with a cumulative mileage of 200740 KM. The data processing flow is shown in Figure 3.

TABLE 5: Membership function parameter.

Membership function parameter	x1	x2	x3	x4
U11	0.1058	0.1531	0.1559	0.2032
U12	0.0076	0.0109	0.0111	0.0145
U13	180.71	227.89	249.48	296.66
U21	0.00089	0.0128	0.0131	0.017
U22	0.1377	0.1647	0.1856	0.2127
U23	0.0087	0.0123	0.0127	0.0163
U31	24.925	32.955	35.172	43.202
U41	0.0586	0.0854	0.0867	0.1135
U42	0.0474	0.0678	0.0694	0.0899
U43	0.0435	0.0634	0.0643	0.0843

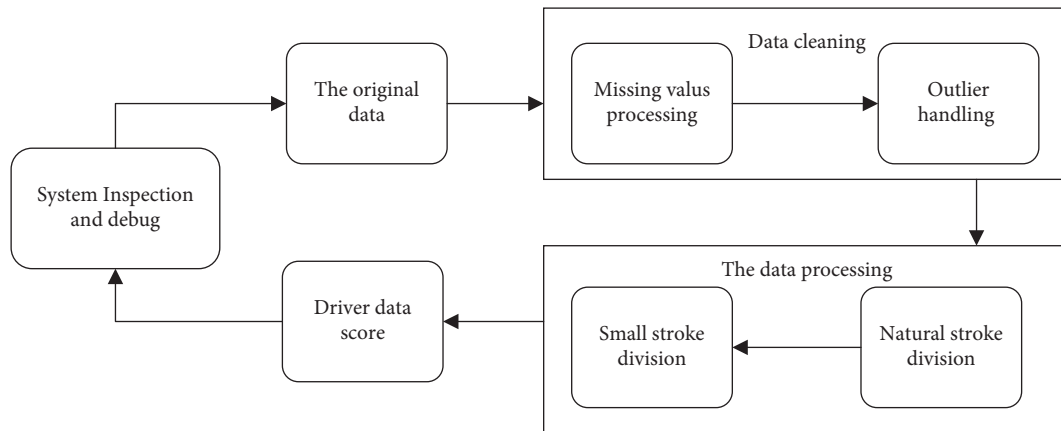


FIGURE 3: Data processing flow chart.

4.2. *Validation of Driving Behavior Economy Evaluation Model and Analysis of Results.* Economic evaluation of driving behavior is conducted within each microstroke segment of the driver, and some representative evaluation results are selected as shown in Figures 4–6. Drivers 5 and 22 represent low frequency of bad driving behavior; drivers 26 and 33 represent medium frequency of bad driving behavior; drivers 3 and 28 represent high frequency of bad driving behavior. The above results indicate that the evaluation system can reflect the difference of drivers’ behavior in different microtravel segments. To further verify the correctness of the evaluation system, check the original index data.

Find the original indicator data based on the low score for verification, as shown in Figure 7. By comparing the reference values of each indicator, it is found that the data of bad driving behavior indicators corresponding to the lower score are more than those exceeding the normal value. This indicates that the evaluation system well reflects the influence of driving behavior on the scoring, and proves the correctness of the evaluation system results.

Further analysis of the results shows that different drivers have different operational performance of bad driving behaviors, as shown in Figure 8. Among them, drivers with serial numbers 2th, 14th, 16th, 29th and 33th all showed a preference for bad driving behaviors with excessive

speed. Therefore, it is necessary to modify the driving habits of these drivers for rapid acceleration.

The drivers with serial numbers 1th, 2th, 11th, 16th, 29th and 33th all have relatively high number of bad driving behaviors, which need to be paid attention to.

In the overall data performance of bad driving behavior, see Figure 9 Different drivers have certain similarities in the performance preference of bad driving behaviors, and the frequency data of bad driving behaviors are further compared. Among them, the indicators with more bad driving behavior frequency are sharp deceleration, low speed, low speed and fast acceleration.

In the overall data performance of bad driving behavior, see Figure 9 Different drivers have certain similarities in the performance preference of bad driving behaviors, and the frequency data of bad driving behaviors are further compared. Among them, the indicators with more bad driving behavior frequency are sharp deceleration, low speed, low speed and fast acceleration.

4.3. *Analysis of Energy Saving Potential.* The driving behavior energy-saving potential of the driver refers to the potential possibility of improving the fuel economy of the vehicle after improving driving behavior. In equations (1) and (2), the relationship between a single driving behavior indicator and

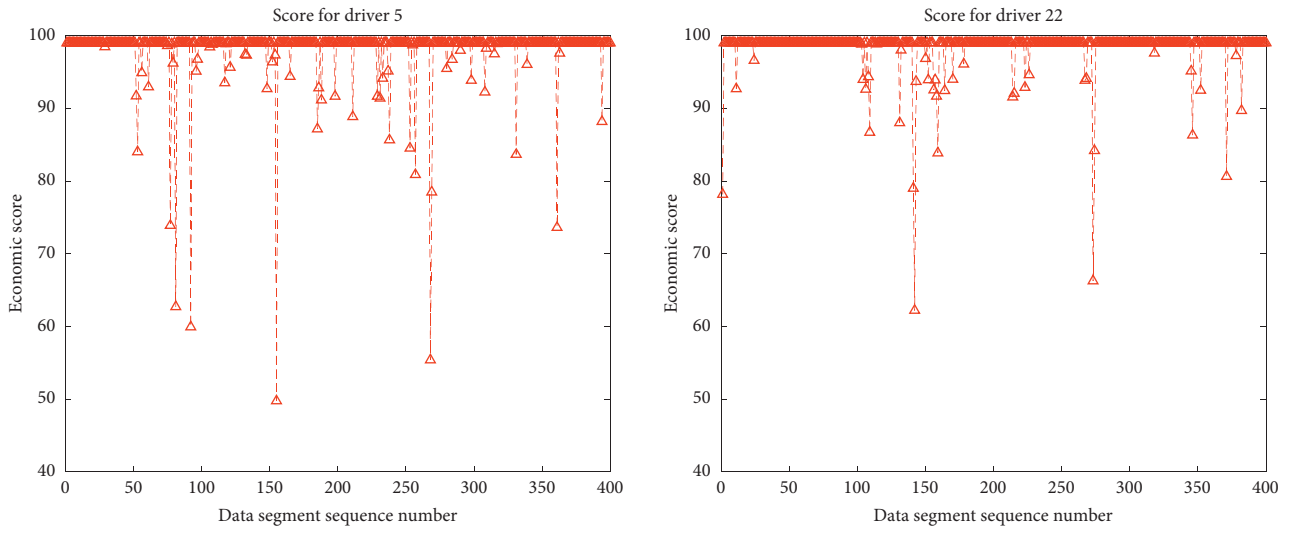


FIGURE 4: Part of the driver's score.

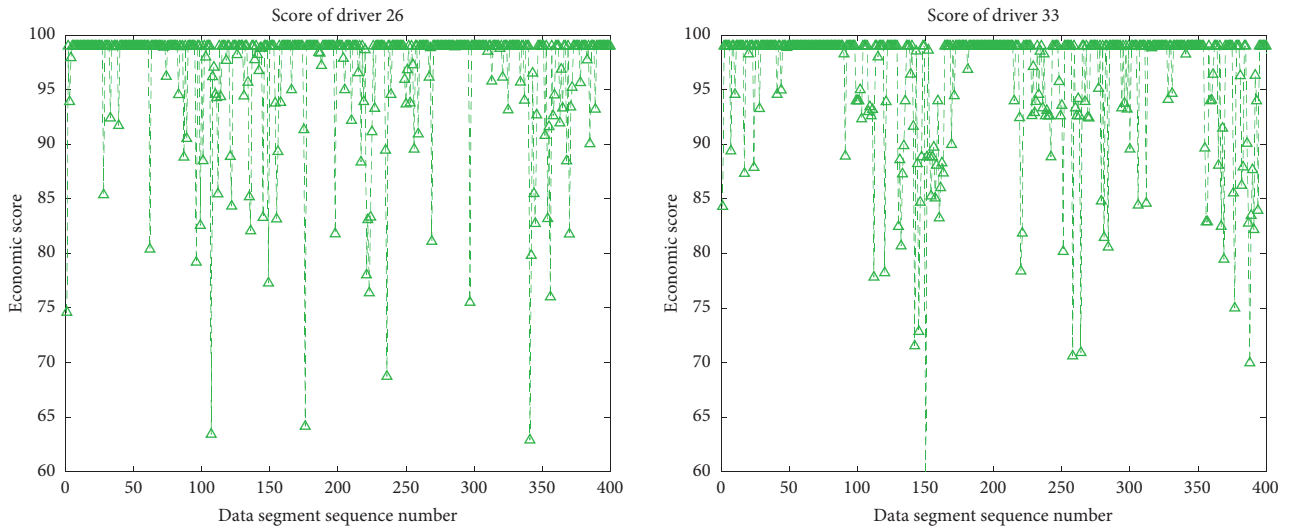


FIGURE 5: Part of the drivers score.

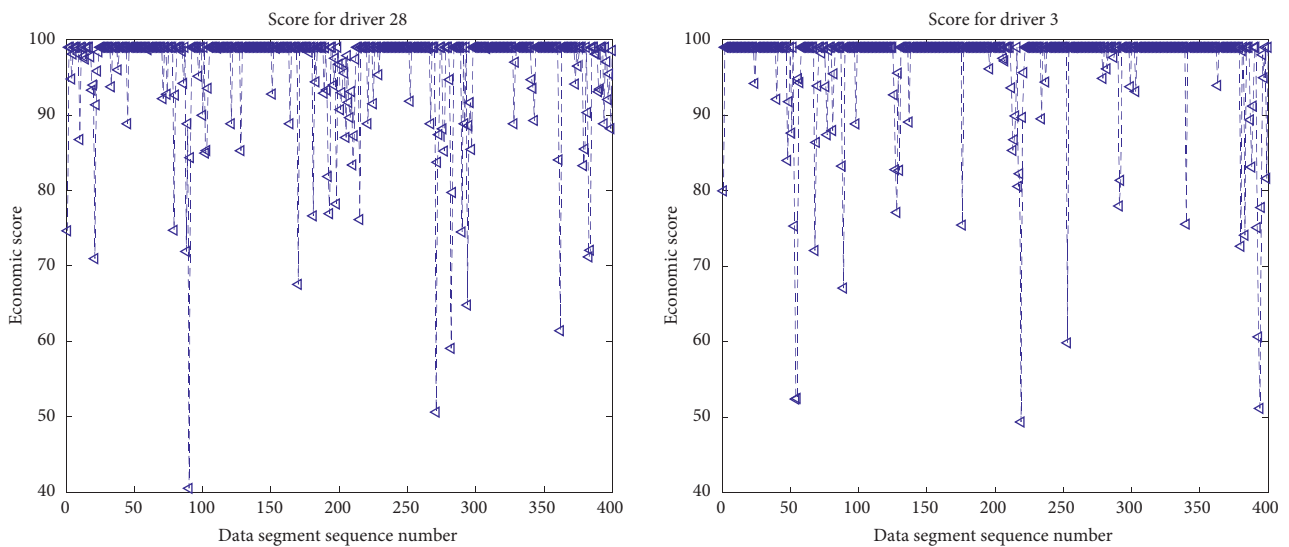


FIGURE 6: Part of the driver's score.

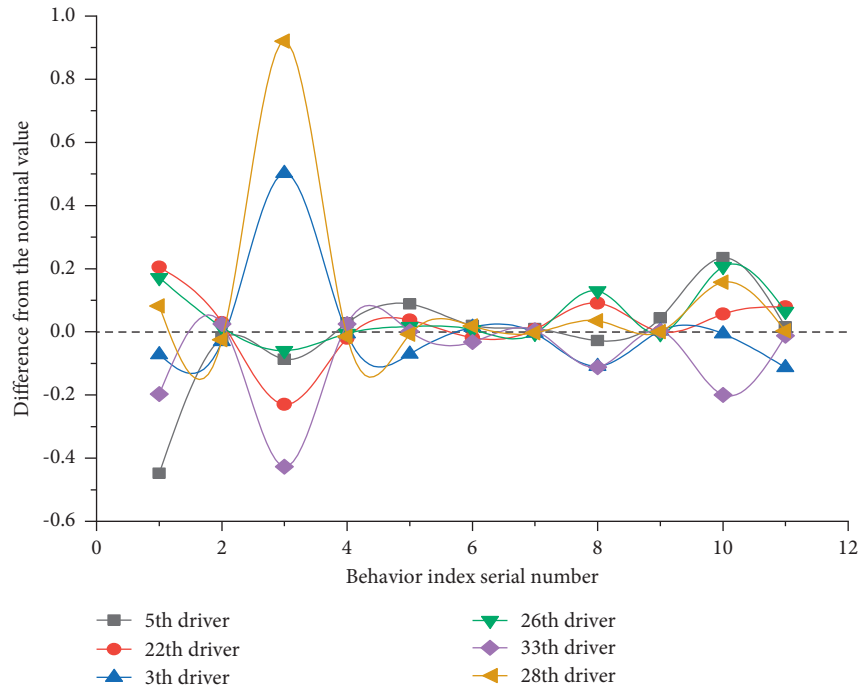


FIGURE 7: Test the low score index data.

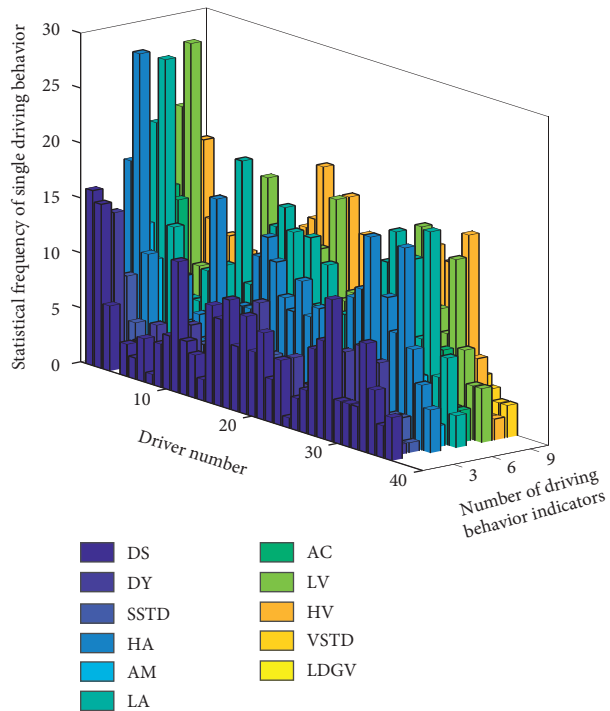


FIGURE 8: The statistical frequency chart of drivers' single driving behavior performance.

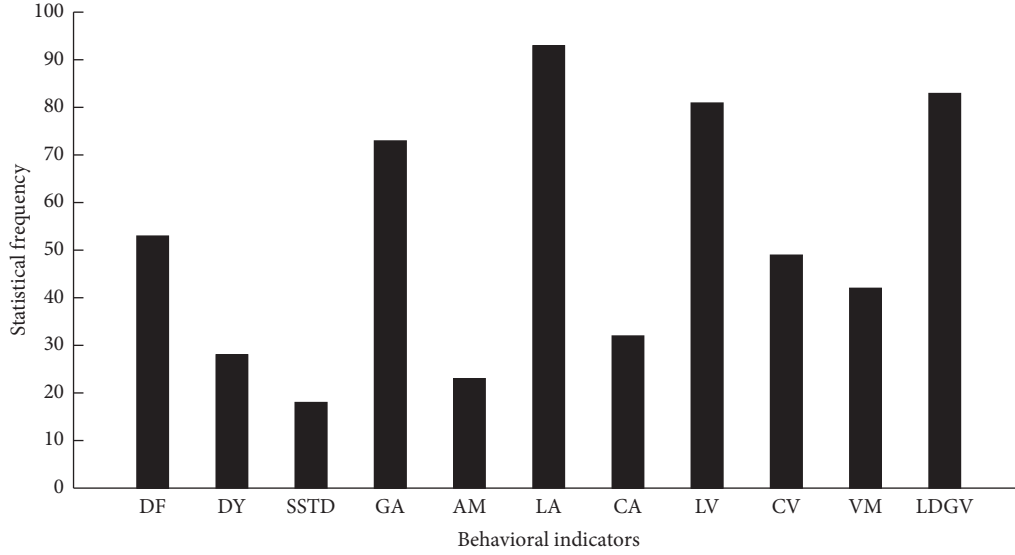


FIGURE 9: Overall frequency of driving behavior indicators.

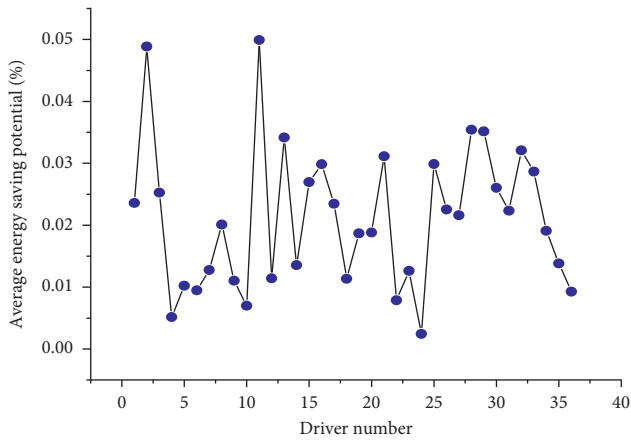


FIGURE 10: Average energy saving potential of drivers.

fuel consumption is given, which is further extended to all driving behavior economic indicators, as shown in equations (21) and (22).

$$F_{\text{loss}} = \sum_{i=1}^m \Delta F_i, \quad (21)$$

$$Q_{\text{loss}} = \sum_{i=1}^m \Delta Q_i. \quad (22)$$

It can be considered that there is a corresponding relationship between driving behavior loss score and driving behavior loss fuel consumption, that is, the improvement of each driving behavior score represents the reduction of fuel consumption and the improvement of fuel economy. Therefore, based on the highest score of driving behavior economy, a calculation method of energy saving potential is given based on the relationship between driving behavior economy score and fuel consumption.

$$\beta = \frac{F_{\text{loss}}}{F_{\text{max}}} \times 100\%. \quad (23)$$

By analyzing and calculating the energy saving potential of different drivers, the results are shown in Figure 10 in which drivers 2 and 11 have higher average energy saving potential compared with other drivers, which has a large space for optimization. The average maximum energy saving potential is 4.98%.

5. Conclusions

In this study, based on correlation analysis and fuzzy comprehensive hierarchical evaluation algorithm, a driving behavior economy evaluation system is established and a calculation method of energy saving potential is proposed. Correlation analysis plays a good role in screening driving behavior indicators. The fuzzy analytic hierarchy process (FAHP) has a strong applicability to the evaluation of driving behavior economy, which is influenced by multiple factors. The evaluation system and the calculation method of energy saving potential are tested by using the data of vehicle Internet of real cars. The results demonstrate the correctness of the work. In addition, the results of further analysis show that the overall driving behavior preference of drivers shows that the frequency of rapid acceleration and deceleration, low speed and low speed behavior is more frequent. The driving behavior preference performance of individual drivers is obviously different; The bad driving behavior of some drivers has a high energy saving potential of up to 4.98%. The study analyzes the differences in bad behaviors in the driver's economic evaluation, which can provide a theoretical basis for the targeted improvement of bad driving behaviors of drivers, and at the same time, the energy conservation analysis provides a basis for the formulation of the subsequent optimization strategies for energy-saving driving behaviors. The research is mainly aimed at freight vehicles, and the relevant system parameters have some

limitations. The follow-up work should further explore the universality of different vehicle types and operating conditions.

Under the background of engineering education, the joint training mode combining theoretical research and engineering practice is more suitable for students to carry out graduation project research and cultivate applied talents for enterprises than the traditional single school training mode, which should be implemented in more university training plans.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the Science and Technology Major Project of Guangxi under Grant AA18242033, Project of Guangxi key Research and Development under Grant AB21196029, Guangxi Natural Science Foundation Program Grant 2020GXNSFAA159081 and 2018GXNSFAA281276, Scientific Research and Technology Development in Liuzhou under Grant 2020GAAA0404, 2021AAA0112 and 2021AAA0104, Guangxi Science and Technology Base and Special Talents Program Grant 2018AD19077, Guangxi Key Laboratory of Manufacture System and Advanced Manufacture Technology under Grant 17-259-05-009Z, Guangxi Higher Education Undergraduate Teaching Reform Project under Grant 2021JGA180, 2020JGA173, and JGY20211085, and Teaching Reform Project of Guilin University of Electronic Technology under Grant JGB202002.

References

- [1] T. J. Daun, D. G. Braun, C. Frank, S. Haug, and M. Lienkamp, "Evaluation of driving behavior and the efficacy of a predictive eco-driving assistance system for heavy commercial vehicles in a driving simulator experiment," in *Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems*, pp. 2379–2386, The Hague, Netherlands, October, 2013.
- [2] R. Burke and C. Brace, "The effects of engine thermal management on performance, emissions and fuel consumption," in *Proceedings of the Sae World Congress & Exhibition*, Detroit, MI, USA, April, 2010.
- [3] H. Liu, L. Han, and Y. Cao, "Improving transmission efficiency and reducing energy consumption with automotive continuously variable transmission: a model prediction comprehensive optimization approach," *Applied Energy*, vol. 274, Article ID 115303, 2020.
- [4] J. C. González Palencia, T. Furubayashi, and T. Nakata, "Energy use and CO₂ emissions reduction potential in passenger car fleet using zero emission vehicles and lightweight materials," *Energy*, vol. 48, no. 1, pp. 548–565, 2012.
- [5] I. De Vlieger, "On board emission and fuel consumption measurement campaign on petrol-driven passenger cars," *Atmospheric Environment*, vol. 31, no. 22, pp. 3753–3761, 1997.
- [6] J. Van Mierlo, G. Maggetto, E. Van de Burgwal, and R. Gense, "Driving style and traffic measures-influence on vehicle emissions and fuel consumption," *Proceedings of the Institution of Mechanical Engineers - Part D: Journal of Automobile Engineering*, vol. 218, no. 1, pp. 43–50, 2004.
- [7] B. Krishnamoorthy and S. Gopalakrishnan, "Truck driver's driving performance assessment," in *Proceedings of the 2008 Commercial Vehicle Engineering Congress & Exhibition*, Chicago, IL, USA, January, 2008.
- [8] R. Frank, G. Castignani, R. Schmitz, and T. Engel, "A novel eco-driving application to reduce energy consumption of electric vehicles," in *Proceedings of the International Conference on Connected Vehicles & Expo IEEE*, Las Vegas, USA, December 2013.
- [9] C. Rolim, P. Baptista, G. Duarte, T. Farias, and Y. Shiftan, "Quantification of the impacts of eco-driving training and real-time feedback on urban buses driver's behaviour," *Transportation Research Procedia*, vol. 3, pp. 70–79, 2014.
- [10] J. C. Ferreira, J. D. Almeida, and A. R. D. Silva, "The impact of driving styles on fuel consumption: a data-warehouse-and-data-mining-based discovery process," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2653–2662, 2017.
- [11] H. Chia-Yu, S. L. Sirirat, and S. Y. Chin, "Data mining for enhanced driving effectiveness: an eco-driving behavior analysis model for better driving decisions," *International Journal of Production Research*, vol. 55, no. 23, pp. 7096–7109, 2017.
- [12] N. H. Tung and T. L. Hoang, "Driving behavior in mixed traffic flow: a novel model for assessing bus movement considering the interaction with motorcyclists," *IATSS Research*, vol. 44, no. 2, pp. 125–131, 2019.
- [13] Y. Saboohi and H. Farzaneh, "Model for developing an eco-driving strategy of a passenger vehicle based on the least fuel consumption," *Applied Energy*, vol. 86, no. 10, pp. 1925–1932, 2008.
- [14] H. D. Gamage and J. Lee, "Machine learning approach for self-learning eco-speed control," in *Proceedings of the Australasian Transport Research Forum Proceedings*, pp. 1–14, Melbourne, Australia, January, 2016.
- [15] K. Boriboonsomsin, A. Vu, and M. Barth, *Environmentally Friendly Driving Feedback Systems Research and Development for Heavy Duty Trucks* Institute of Transportation Studies., Davis, GA, USA, 2016.
- [16] I. Jeffreys, G. Graves, and M. Roth, "Evaluation of eco-driving training for vehicle fuel use and emission reduction: a case study in Australia," *Transportation Research Part D*, vol. 60, pp. 85–91, 2016.
- [17] J. Díaz-Ramírez, N. Giraldo-Peralta, D. Flórez-Ceron, V. Rangel, and C. M. Argueta, "Eco-driving key factors that influence fuel consumption in heavy-truck fleets: a Colombian case," *Transportation Research Part D: Transport and Environment*, vol. 56, pp. 258–270, 2017.
- [18] H. J. Ma, H. Xie, and D. Brown, "Eco-driving assistance system for a manual transmission bus based on machine learning," *IEEE Transactions on Intelligent Transportation*, vol. 19, no. 2, pp. 572–581, 2017.
- [19] C. Chen, X. Zhao, Y. Yao, Y. Zhang, J. Rong, and X. Liu, "Driver's eco-driving behavior evaluation modeling based on

- driving events,” *Journal of Advanced Transportation*, vol. 2018, Article ID 9530470, 12 pages, 2018.
- [20] R. Hao, H. Yang, and Z. Zhou, “Driving behavior evaluation model Base on big data from internet of vehicles,” *International Journal of Ambient Computing and Intelligence*, vol. 10, no. 4, pp. 78–95, 2019.
- [21] Z. Xu, T. Wei, S. Easa, X. Zhao, and X. Qu, “Modeling relationship between truck fuel consumption and driving behavior using data from internet of vehicles,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 3, pp. 209–219, 2018.
- [22] H. Zheng, Y. Wu, Z. Wang, and Z. Zhang, “AHP based driving behavior evaluation model,” *Journal of Physics: Conference Series*, vol. 1325, 2019.
- [23] Z. W. Qian, Y. F. Shi, and X. L. Du, “Selection of logistics service provider based on fuzzy comprehensive evaluation,” in *Proceedings of the 2016 International Conference on Management Science and Management Innovation*, Guilin, China, August, 2016.
- [24] A. Bigand and O. Colot, “Membership function construction for interval-valued fuzzy sets with application to Gaussian noise reduction,” *Fuzzy Sets and Systems*, vol. 286, pp. 66–85, 2015.

Research Article

Enhancing Problem-Solving Ability through a Puzzle-Type Logical Thinking Game

Ting-Sheng Weng 

Department of Business Administration, National Chiayi University, Chiayi 600, Taiwan

Correspondence should be addressed to Ting-Sheng Weng; politeweng@yahoo.com.tw

Received 28 October 2021; Revised 23 December 2021; Accepted 18 February 2022; Published 24 March 2022

Academic Editor: Wenbing Zhao

Copyright © 2022 Ting-Sheng Weng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study explored the influence of creativity and learning attitude through game-based learning. The subjects were sixth grade elementary school students, and a logical thinking game was developed to stimulate the students' learning interests. A questionnaire survey was conducted on 121 students, and the result shows that the students' creativity has a significant and positive effect on their learning attitude and problem-solving ability. In addition, learning attitude has a significant and positive effect on the students' problem-solving ability. Moreover, creativity elicited a direct effect on problem-solving ability. This study verified that the proposed puzzle-type logical thinking game can train and improve the logical thinking ability of learners.

1. Introduction

Problem-solving ability has been valued in education. Present educational approaches have encouraged learning through practical problems or cases, inspiring learners to demonstrate creative thinking, and designing solutions to problems. Actual life problems or difficulties are optimal learning examples [1]. The concept of STEM (science, technology, engineering, and mathematics) has gradually been valued and applied in engineering education in colleges and universities [2, 3].

Prensky named the learners in the recent decade as the "digital natives," meaning that they are accustomed to using various technological products in their lives. Games have an open point of view and can bring people the desire to create and invest, and such elements have motivated people's willingness to take on risks and explore [4]. According to Prensky [5], the ideal digital game-based learning is a method to increase learner engagement and generate higher learning achievements. To improve problem-solving ability, learners' creativity must first be stimulated, and game-based teaching is a feasible approach to create this stimulation. Using games to present the teaching contents not only makes the contents more vivid but also inspires the imagination of learners and triggers more creativity. Thus, game-

based teaching is an effective method for eliciting creative thinking [6–10]. Li [11] used tangram puzzles as a representation of mathematics and posited that it could stimulate game-based learning intention. Knowledge can be internalized in the process of puzzle playing, and speed and skill refinement can also be generated during the game-playing process. Gao [12] stated that games can stimulate learners' enthusiasm for learning and inspire their initiative, independence, and creativity. Through playing games, learners can immerse themselves in those activities and are encouraged to think about, understand, digest, and absorb details from concrete imaginary thinking to abstract thinking. Therefore, learning interest and knowledge grow naturally in a relaxed and pleasant atmosphere. Furthermore, game-based learning also helps to cultivate learners' thinking ability and quality, learning attitude, awareness and methods of observing surrounding objects, inferential logic, exploration spirit, and ability to think outside the box to achieve logic creation.

However, conventional educational toys that exist as physical entities have no characterization orientation, as their representational state is fixed. Digital learning appears to be the solution to this problem, as it uses multimedia applications to create a variety of educational platforms (e.g., video images, dynamic images, and model simulations), in

which teaching materials can be presented in various characterizations that enrich the learning process. In game-based teaching, multimedia special effects and interactive characteristics can attract learners' attention, thereby enhancing their learning motivation and improving their problem-solving abilities [13, 14]. Through game-based practices in the context of digital games, learners can be guided to become "experiencers" and personally solve multilevel problems created in the games. The creativity and problem-solving ability that the learners accumulated could lead them to become knowledge creators that provide solutions to real-world problems [15].

It is evident that teaching content and educational methods can stimulate learners' creativity, problem-solving ability, and solution development efficacy [16–20]. Learners' creativity has a direct effect on their problem-solving ability and may affect learning attitude, which has an indirect effect on their problem-solving ability.

Therefore, this study suggests that the use of multimedia technology is a viable approach to developing a more interactive and stimulating puzzle-type logical thinking game that allows learners to engage in creative thinking, thereby affecting their learning attitude and improving their problem-solving ability. This study aims to explore the influences of digital learning tools on learners' learning attitude, problem-solving ability, and creativity. More specifically, this study intends to verify whether game-based learning can enhance learners' creative thinking and further influence their learning attitude and generate positive effects on their problem-solving ability.

2. Literature Review

2.1. Game and Learning Attitude. Many learners hold a passive attitude toward learning. Attitude toward science refers to whether a person likes or dislikes science or has "a positive or negative feeling about science" [16]. Xu and Liao [21] used mathematics learning as an example to demonstrate that students may be uninterested in learning mathematics due to learning disabilities, poor language ability, and difficulties in understanding abstract symbols. They suggested that image interface guidance in games could help learners to understand mathematical logic concepts. Su and Xie [22] mentioned that game-based teaching activities can easily inspire learners' explorative ability toward a certain subject, enhance their learning attitude, attract their attention to problems, and improve their problem-solving ability. Zheng and Liu [17] proposed that mission-based games can promote cooperation and discussion among peers. Compared with conventional paper-based tests, learners' learning motivation, attitude, and achievements were substantially improved after playing mission-based games. Yang and Li [19] experimented with a mathematics game and found that the absolute use of technology can be a helpful learning tool for developing children's number sense [19].

It is very important to create an environment that can promote and maintain learners' interest [23]. Hwang et al. [24] pointed out that players' attitudes affect their subsequent use of the game, and players attach great importance

to their game experiences [25, 26]. The desire to overcome obstacles in the game experience will encourage players to find ways to achieve breakthroughs and solve problems, which in turn improves their problem-solving ability. As mentioned above, the teaching methods that employ puzzle-type logical thinking games can improve learners' learning motivation and attitude, enable learners to focus on how to complete a challenge, and stimulate the brain's hidden creativity during a short play time. In the process of attempting to complete a challenge, learners tend to brainstorm and identify the optimal approaches to succeed by repeatedly reviewing and evaluating the chosen approach, identifying problems, and overcoming obstacles. Finally, positive learning experiences can be accumulated from game-based learning.

It is effective to learn subjects through games with logical thinking factors. For example, Sun-Lin and Chiou [27] used a game-based learning strategy to support algebra variable learning and found that their proposed method could effectively enhance students' learning methods and learning attitudes. Hwa [28] took the mathematics classes of first to third grade students in Chinese elementary schools as subjects and found that game-based learning is superior to the traditional classroom instruction in terms of students' logical thinking, learning attitude, and optimizing the teaching process.

2.2. Creativity. People that think creatively tend to confidently put forward different opinions and try solutions that are different from the status quo. They may combine the knowledge of different fields to solve problems and persist instead of giving up in times of difficulties and no progress. Moreover, they may even deliberately set aside a problem temporarily and return to the problem later after searching for the solution from a new perspective [29–31]. Lee and Therriault [32] suggested that creativity is a high-level cognitive activity generated by idea association, as well as divergence and convergence of thought by individuals, through the collection of data and memory. Gilhooly et al. [33] also posited that creativity and thinking mode are correlated. Xiao [34] proposed that improvements in learners' creativity can be achieved through intrinsic motivation, teaching content, and the reading atmosphere in the learning environment. In terms of teaching, open-ended questions and extended questions can encourage learners to think repeatedly and enhance their creativity. Furthermore, Xiao [34] stated that creativity includes the aspects of adventure, curiosity, imagination, and challenge. Xiao further suggested that creativity is inspired and enhanced by parental creativity education, creativity life experience, and daily life experience.

In addition, teachers' instructional strategies can inspire students' creativity. Yang et al. [35] designed the educational strategy of robot-assisted instruction to support students' learning achievements. The results showed that an appropriately designed teaching strategy can enable students to think and recognize learning materials through the interactive models.

Based on the definitions in previous literature, this study defined creativity as follows. Teachers use open teaching materials to guide learners to think independently and inspire learners to develop new ideas and actions outside of in-class assignments; by helping the learners to probe into the root of problems and relevant causes, teachers can guide learners to consider possible approaches to solving these problems.

2.3. Game-Based Learning and Problem-Solving Ability. In problem solving, the process is equally important as the outcome [36]. Polya [37] suggested that problem solving is an explicit or cognitive behavioral process in which various possible effective solutions are proposed for the problematic situation, and individuals select the most effective solution for implementation. Liao and Wang [38] stated that in problem solving, the thought process is crucial. They summarized the psychological process as problem identification, understanding the nature of the problem, collecting relevant information, taking problem-solving actions, and engaging in postevaluation. The “problem” in their study was the gap between the goal and current situation, and problem solving is a process in which an individual applies one’s knowledge to the problem to obtain solutions. Lin and Huang [18] argued that problem-solving attitude is the basis for problem-solving ability. They posited that problem-solving attitude and ability affect individuals’ thoughts and motivations when dealing with problems.

Nevertheless, conventional education models have a limited effect on the cultivation of problem-solving ability. Most learners are accustomed to relying on teachers’ guidance and lack the ability to think independently. Therefore, appropriate teaching activities could cultivate learners’ creativity and problem-solving ability [39]. Yang [40] stated that receptive or inquiry-based teaching is not the only valid teaching method, but it should be conducted simultaneously. The goal of teaching is not merely imparting knowledge but to help learners understand thinking and the thought process. Teaching should not only be for imparting knowledge but also to help learners understand thinking and the thought process; when knowledge concepts are understood and appropriately applied, they can be effectively used to help learners solve problems and cultivate their thinking ability.

Tang [20] posited that competition in society is becoming more intense, and that while problem-solving ability has gradually been recognized and valued, it has become an indispensable ability for talented people. Nevertheless, traditional education models have reduced learners’ interest in learning and have failed to improve their problem-solving ability. Tang stated that developing a game-based learning model is imperative to improving learners’ problem-solving ability. Huang et al. [41] indicated that game scores are significantly correlated with learners’ graphic reasoning, creative thinking, and problem-solving abilities. Furthermore, game scores are related to learners’ higher-order thinking ability. Therefore, games are suitable tools for training learners’ creativity and problem-solving ability. Zheng [42] stated that mathematics-based games in elementary schools could train learners’ problem-solving ability.

Many related studies have explored learning outcomes from game-based learning integrated with education in the field of digital learning [43–46], with system development [47] as the research goal. This study investigated whether game-based learning has a positive effect on learners’ problem-solving ability by exploring the creativity learning attitude and the psychological level of learners.

With highly developed multimedia technology, creativity educational materials with sound and light effects and visual animation can be designed. An interactive interface with visual elements, such as images, can be designed to meet the learners’ needs. Liao and Wang [38] designed an exploratory game and added an element of fun to the learning process. Their game produced excellent learning outcomes and significant improvement in the subjects’ problem-solving ability. Moreover, they demonstrated that game scenarios can influence problem-solving methods and cultivate learners’ creativity and thinking from different perspectives. Game-based learning activities can enhance learning motivation and inspire learners to apply their creative thinking skills in a team setting to solve problems. By adopting game-based teaching activities, novel teaching methods or various highly interactive teaching models can be developed [48, 49].

Among various digital games, computer games have attracted the attention of young students, as well as children. Game-based education is still effective in guiding and inspiring children’s judgmental thinking abilities. There are numerous research studies exploring the effectiveness of game-based learning and many curricula adopting learning games [50–54]. In particular, puzzle-type logical thinking games can lead learners to engage in creative thinking and develop problem-solving ability.

3. Research Method

This study used Flash software to develop a puzzle-type logical thinking game on the website (Figures 1 and 2). Figure 3 presents a schematic of a stroke-completing game as a test tool for a creative thinking activity. The research tool involved a puzzle game that has not been frequently applied in general teaching, and it was designed to stimulate young students’ curiosity and interest in learning. This study took sixth grade elementary school students as the subjects. The questionnaire survey method was adopted to measure the subjects’ creativity, learning attitude, and attitude toward game-based learning and problem solving. The data were collected based on the four dimensions of creativity, learning attitude, attitude toward game-based learning, and attitude toward problem solving.

3.1. Research Scope and Participants. This study was based on a puzzle-type logical thinking game. Purposive sampling was used to select learners who participated in science exhibition activities. Among the 150 participants, there were 121 valid participants, with a valid response rate of 80.6%.

3.2. Research Tools. The research structure of this study is presented in Figure 4. A linearly structured model was used to establish the relationship between and different influences

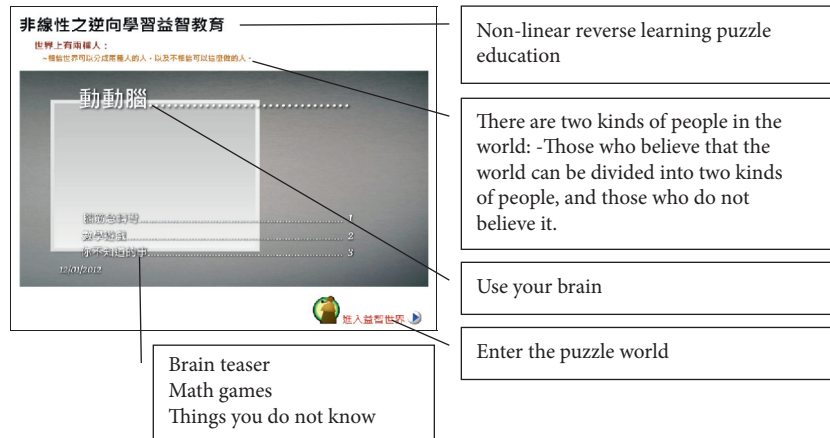


FIGURE 1: Puzzle game portal.

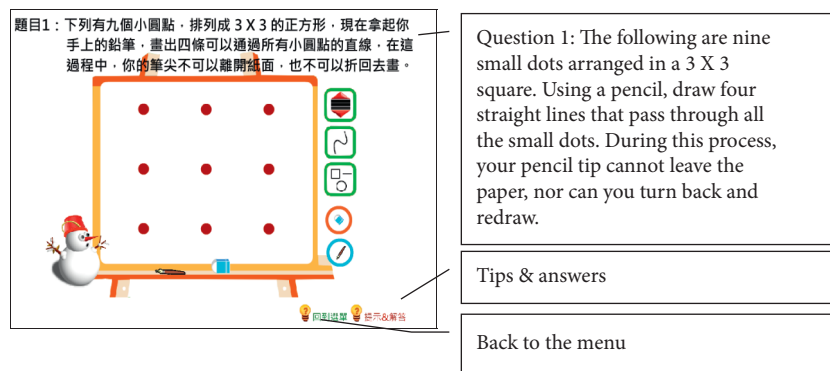


FIGURE 2: Screen of “thinking after corner-turning” puzzle game.

of the four dimensions of creativity, learning attitude, attitude toward game-based learning, and attitude toward problem solving. The items for the attitude toward game-based learning, learning, creativity, and problem solving were modified from the questionnaires of Chu et al. [55]; Chen et al. [56]; Cao and Zhou [57]; Amabile [29]; Zhou and George [58]; and Lin and Huang [18], respectively. For this questionnaire survey, a 5-point Likert scale was used for measuring the three dimensions of creativity, learning attitude, and problem solved.

There were 121 valid samples in this study, which is likely to deviate from the assumption of normal distribution. Compared to AMOS, which assumes normal distribution based on a large sample, SmartPLS uses a PLS-SEM model that supports the assumption of abnormal distribution. Therefore, this study used SmartPLS M2.0.3 for statistical analyses. The operational definitions of the four dimensions are shown in Table 1.

3.3. Experiment Procedure. The experimental procedure is as follows: (1) the researchers explained the game to the subjects; (2) the subjects played the game; (3) questionnaires were distributed; (4) questionnaires were collected; and (5) data analyses were performed.

For data analysis, structural equation modeling was employed to measure the directions of influence among the

four dimensions. Reliability and validity analyses were used to examine whether the questionnaires used in this study was suitable for analysis, and effective results were obtained.

3.4. Reliability and Validity Analyses. The scales used in this study were modified from existing studies. Thus, Cronbach's α value of each dimension ranged from 0.697 to 0.926, which satisfied the reliability of index requirements. Moreover, the scales were also modified based on the expert opinions, confirming that the scales have satisfactory expert validity (Table 2).

4. Results

A linearly structured model was applied to determine the relationships and influences between the four dimensions. Furthermore, the SmartPLS M2.0.3 software suite was used to test the relationship between the four dimensions and goodness of fit of the model. The linear model is shown in Figure 5. The overall model goodness of fit was determined using the overall goodness of fit of validation mode and observation data. A higher goodness of fit of a model indicates that it is highly feasible. The results of this study indicate that most of the overall model indices fitted well with ideal values (Table 3); thus, the goodness of fit of the model was within an acceptable range.

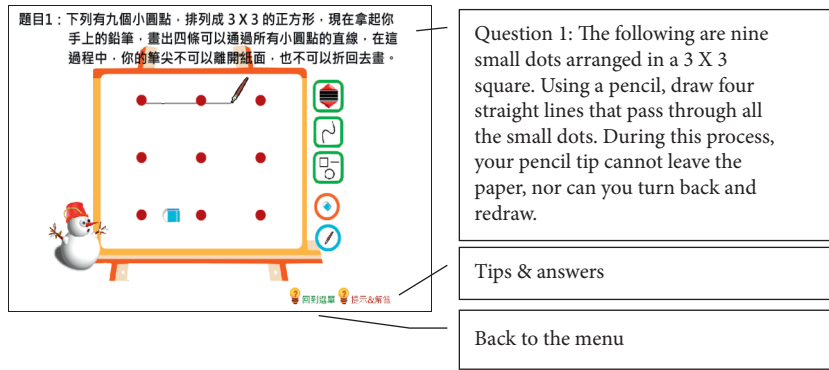


FIGURE 3: Stroke movement.

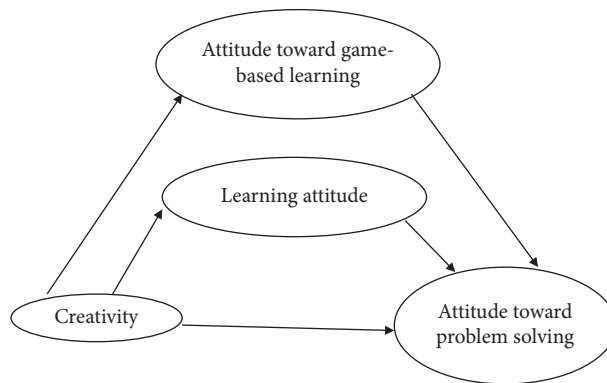


FIGURE 4: Research model.

TABLE 1: Dimensions, definitions, and citations.

Dimensions	Definitions	References
Creativity	Students evaluate whether they are producing answers in classwork that can be used to solve assignments because of the tool	Lee and Therriault [32]; Gilhooly et al. [33]; Yang et al. [35] Koballa and Crawley [16];
Learning attitude	Students' cognition and values of "learning"	Su and Xie [22]; Garcia et al. [23]; Xu and Liao [21]; Yang and Li [19]; Hwang et al. [24] Richard [48]; Huang et al. [41]; Lee and Lee [47]; Tang et al. [49]; Zheng [42];
Attitude toward game-based learning	Students' cognition and evaluation of learning content or learning activities using this research tool	Halloluwa et al. [43]; Muñoz et al. [45]; Vrugte et al. [46]; Lee et al. [44] Polya [37]; Li [39]; Yang [40];
Attitude toward problem solving	Students' cognition and values of "problem solving"	Liao and Wang [38]; Lin and Huang [18]; Tang [20]

TABLE 2: Reliability and validity analyses.

Dimension	Cronbach's α
Creativity	0.917
Learning attitude	0.717
Attitude toward game-based learning	0.697
Attitude toward problem solving	0.926

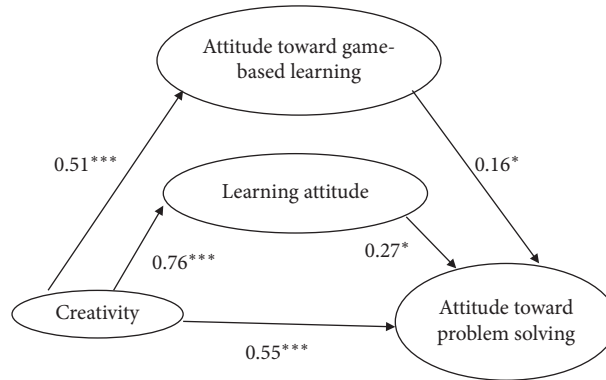


FIGURE 5: Results of research model.

TABLE 3: Goodness of fit of the overall model.

Adaptation index	χ^2 test	χ^2/df	GFI	AGFI	PGFI	RMSEA	NNFI	CFI
Suggested values	$p > 0.05$	< 3.0	> 0.9	> 0.9	> 0.5	< 0.08	> 0.9	> 0.9
Research model	$\chi^2 = 2131.869, p \leq 0.001$	1.991	0.823	0.678	0.622	0.092	0.566	0.719

The relationships between creativity, learning attitude, attitude toward game-based learning, and problem-solving are described as follows (Figure 2 and Table 4):

- (1) Learner's creativity has a significant and positive effect on learning attitude, attitude toward game-based learning, and problem solving ($\beta_1 = 0.51, t = 3.955; \beta_2 = 0.76, t = 6.201; \beta_3 = 0.55, t = 4.017$).
- (2) Learning attitude and attitude toward game-based learning have a significant and positive effect on problem solving ($\beta_4 = 0.16, t = 2.022; \beta_5 = 0.27, t = 2.309$).
- (3) Learners' creativity has a direct effect, an indirect effect, and a total effect of 0.55, 0.2868, and 0.8368, respectively, on problem solving.

According to the analysis results, creativity, learning attitude, and attitude toward game-based learning affect problem solving. This finding is consistent with the setting of the original research (Figure 1), suggesting that learning attitude, attitude toward game-based learning, and creativity have a significant effect on the ability of learners to solve problems while faced with difficulties. The results of the linear model analysis indicate that creativity also affects learning attitude and attitude toward game-based learning. This result implies that learners' willingness to accept new things may influence their learning attitude and attitude toward game-based learning (Table 4).

TABLE 4: Path coefficients affecting the dimensions.

Path		Standardized coefficients	p values
Learning attitude	<--- Creativity (β_1)	0.51***	≤ 0.001
Attitude toward game-based learning	<--- Creativity (β_2)	0.76***	≤ 0.001
Attitude toward problem solving	<--- Learning attitude (β_4)	0.16*	0.043
Attitude toward problem solving	<--- Attitude toward game-based learning (β_5)	0.27*	0.021
Attitude toward problem solving	<--- Creativity (β_3)	0.55***	≤ 0.001

*** $p < .001$; * $p < .05$; ** $p < .01$.

This study explored the relationships among learning attitude, attitude toward game-based learning, and creativity regarding problem solving. The results indicate that creativity, learning attitude, and attitude toward game-based learning have significant effects on problem solving. The findings of this study can be further explored by referring to

digital learning. This type of learning can be improved by integrating game-based learning into problem-solving tasks.

5. Conclusion and Suggestions

Computer-aided instruction not only improves problem-solving ability but also learning attitude [59, 60]. This study found that for learners and the general public, learning attitude and attitude toward game-based learning considerably affect their problem-solving ability. In addition, the results of this study confirmed that creativity and learning attitude are crucial and are affected by learners' willingness to learn and the actions taken.

This study treated sixth grade elementary school students as the subjects. Future studies can conduct surveys on learners in different countries. Moreover, qualitative interviews can be conducted to analyze the problem-solving process of the learners and the factors affecting changes in their attitudes toward game-based learning. After these results are obtained, more variables can be generated to identify factors that may influence learners' problem-solving ability. The analysis results may allow teachers to identify optimal learning methods that could enhance learners' problem-solving ability. According to the results of this study, creativity, learning attitude, and attitude toward game-based learning strongly affect problem-solving ability.

The rules of games often lead to constraints that limit thinking within a prescribed range, like a bird in a cage. When looking at certain things and thinking about certain problems, individuals are susceptible to being restrained by traditional thinking and being unable to break through and innovate.

Whether in life or education, it is very important to have the ability to solve problems in the environment, to break game rules to find solutions. Thus, it is recommended that researchers can design more puzzle-type learning games in the future for children or educational units to train students' ability to break through habitual thinking and develop adaptability as an instinct.

The contribution of this study is to demonstrate that game-based learning concept can serve as a basis for remedial education or course management and as reference for curriculum design and interactive models.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] D. S. Hu, "Why do people need to learn mathematics—philosophical thinking of mathematical meaning," *Journal of Mathematics Education*, vol. 4, pp. 54–57, 2010.
- [2] C. M. Hung, G. J. Hwang, and I. Huang, "A project-based digital storytelling approach for improving learners' learning motivation, problem solving competence and learning achievement," *Journal of Educational Technology & Society*, vol. 15, no. 4, p. 368, 2012.
- [3] M. Taub, R. Azevedo, A. E. Bradbury, G. C. Millar, and J. Lester, "Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment," *Learning and Instruction*, vol. 54, pp. 93–103, 2018.
- [4] L. E. Hicks, "Infinite and finite games: play and visual culture," *Studies in Art Education*, vol. 45, no. 4, pp. 285–297, 2004.
- [5] M. Prensky, "Digital natives, digital immigrants part 1," *On the Horizon*, vol. 9, no. 5, pp. 1–6, 2001.
- [6] J. H. Chen, "The Learning Activities of Mathematics to Fuse Both of Plays and Toys," National Science Council Project, NSC94-2515-S-168-001, Taipei, Taiwan, pp. 1–21, 2006.
- [7] X. Lin, *Smart Kid's Favorite Puzzle Game: Brain Teaser*, Chemical Industry Press, Beijing, China, 2014.
- [8] A. All, E. P. Nuñez Castellar, and J. Van Looy, "Assessing the effectiveness of digital game-based learning: best practices," *Computers & Education*, vol. 1, pp. 90–103, 2016.
- [9] C. J. Lin, G. J. Hwang, Q. K. Fu, and J. F. Chen, "A flipped contextual game-based learning approach to enhancing EFL learners' English business writing performance and reflective behaviors," *Journal of Educational Technology & Society*, vol. 21, no. 3, pp. 117–131, 2018.
- [10] E. Yukselturk, S. Altrok, and Z. Başer, "Using game-based learning with kinect technology in foreign language education course," *Journal of Educational Technology & Society*, vol. 21, no. 3, pp. 159–173, 2018.
- [11] Y. Li, "Interesting tangram puzzle teaching games," *Mathematics Teaching*, vol. 2, pp. 8–11, 2015.
- [12] J. Gao, "Correlation analysis of influence of mathematical games on creativity thinking," *Course Education Research*, vol. 12, pp. 257–258, 2015.
- [13] T. Y. Chuang and W. F. Chen, "Effect of computer-based video games on children: an experimental study," *Educational Technology & Society*, vol. 12, no. 2, pp. 1–10, 2009.
- [14] J. W. Fan and X. Z. Zhou, "Research on the integration of computer animation into the teaching design of mathematics at primary school," *New Horizons*, vol. 58, no. 3, pp. 133–148, 2010.
- [15] M. F. Zhan, "Fostering learning paradigm shift with game-based learning," *Journal of Advanced Technology and Management*, vol. 1, no. 1, pp. 47–60, 2011.
- [16] T. R. Koballa and F. E. Crawley, "The influence of attitude on science teaching and learning," *School Science & Mathematics*, vol. 85, no. 3, pp. 222–232, 1985.
- [17] Q. P. Zheng and W. Y. Liu, "A study on the compared effects of blockade-running assessments and paper-and-pencil tests on junior high school learners' math learning," *Science Education Monthly*, vol. 338, pp. 16–28, 2011.
- [18] S. H. Lin and Y. C. Huang, "Relationship between academic frustration tolerance and problem-solving attitude of college learners in Central Taiwan," *Global Mental Health E-Journal*, vol. 2, pp. 25–44, 2011.
- [19] D. C. Yang and M. N. Li, "Assessment of animated self-directed learning activities modules for children's number sense development," *Educational Technology & Society*, vol. 16, no. 3, pp. 44–58, 2013.
- [20] G. Y. Tang, "Primary school mathematics education game designed for problem-solving ability training," *Reading and Write Periodical*, vol. 13, 2015.
- [21] Y. X. Xu and G. Z. Liao, "Investigation on graphic thinking process of mathematics addition games for underachievement

- elementary school learners,” *International Journal on Digital Learning Technology*, vol. 4, no. 1, pp. 17–41, 2012.
- [22] X. L. Su and X. Y. Xie, “The study of science game incorporated into elementary school natural science and technology instructions on problem-solving ability,” *Journal of Scientific and Technological Studies*, vol. 40, no. 1, pp. 47–68, 2006.
- [23] R. M. C. Garcia, C. D. Kloos, and M. C. Gil, “Game based spelling learning,” in *Proceedings of the In 38th Annual Frontiers in Education Conference*, Saratoga Springs, NY, USA, October 2008.
- [24] M.-Y. Hwang, J.-C. Hong, H.-Y. Cheng, Y.-C. Peng, and N.-C. Wu, “Gender differences in cognitive load and competition anxiety affect 6th grade students’ attitude toward playing and intention to play at a sequential or synchronous game,” *Computers & Education*, vol. 60, no. 1, pp. 254–263, 2013.
- [25] J. C. Yen, “A study on the experience value, consumer motivation and repurchase intention of PC online games,” *Tamsui Oxford Journal of Tourism*, vol. 15, pp. 1–16, 2019.
- [26] J. J. Lo and W. J. Kuo, “Impact of personality type and time pressure on players’ gameplay experiences, attitudes, and performance,” *Research of Educational Communications and Technology*, vol. 124, pp. 53–68, 2020.
- [27] H. Z. Sun-Lin and G. F. Chiou, “Effects of comparison and game-challenge on sixth graders’ algebra variable learning achievement, learning attitude, and meta-cognitive awareness,” *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 6, pp. 2627–2644, 2017.
- [28] S. P. Hwa, “Pedagogical change in mathematics learning: harnessing the power of digital game-based learning,” *Journal of Educational Technology & Society*, vol. 21, no. 4, pp. 259–276, 2018.
- [29] T. M. Amabile, “The social psychology of creativity: a componential conceptualization,” *Journal of Personality and Social Psychology*, vol. 45, no. 2, pp. 357–376, 1983.
- [30] M. A. Collins and T. M. Amabile, “Motivation and creativity,” in *In Handbook of Creativity*, R. J. Sternberg, Ed., Cambridge University Press, Cambridge, UK, 1999.
- [31] S. K. Yin, C. C. Yang, and F. R. Liu, “Creativity Design of Serious Games by Students of Engineering and Design – Subproject 1: Curriculum Development of Planning and Marketing,” pp. 1–32, Ministry of Science and Technology, Taipei, Taiwan, 2005, MOST103-2511-S230-002-MY3.
- [32] C. S. Lee and D. J. Theriault, “The cognitive underpinnings of creativity thought: a latent variable analysis exploring the roles of intelligence and working memory in three creativity thinking processes,” *Intelligence*, vol. 41, no. 5, pp. 306–320, 2013.
- [33] K. J. Gilhooly, L. J. Ball, and L. Macchi, “Insight and creative thinking processes: Routineandspecial,” *Thinking & Reasoning*, vol. 21, no. 1, pp. 1–4, 2015.
- [34] J. C. Xiao, “The relationship between schools’ climate of creativity, teachers’ intrinsic motivation, and teachers’ creativity teaching performance: a discussion of multilevel moderated mediation,” *Contemporary Educational Research Quarterly*, vol. 19, no. 4, pp. 85–125, 2011.
- [35] Y. Yang, Y. Long, D. Sun, J. Van Aalst, and S. Cheng, “Fostering students’ creativity via educational robotics: an investigation of teachers’ pedagogical practices based on teacher interviews,” *British Journal of Educational Technology*, vol. 51, no. 5, pp. 1826–1842, 2020.
- [36] E. D. Gagne, C. W. Yekovich, and F. R. Yekovich, *The Cognitive Psychology of School Learning*, Harper Collins, New York, NY, USA, 2nd edition, 1993.
- [37] G. Polya, *Mathematical Discovery: On Understand Learning, and Teaching Problem Solving*, Scribner’s, New York, NY, USA, 1981.
- [38] G. Z. Liao and D. M. Wang, “The Design and Effectiveness Study of Using Creativity Gear Game with Touch Interfaces for Game-Based Learning Problem Solving Ability,” National Science Council Project, NSC 99-2410-H-134-015-MY2, , pp. 1–80, 2012.
- [39] L. Y. Li, “Strategies for cultivating problem awareness in science teaching in junior middle school,” *Journal of Shanghai Educational Research*, vol. 5, pp. 81–85, 2011.
- [40] Q. X. Yang, “Discussion on “integration of numbers and shapes” in primary school mathematics education,” *Forum on Contemporary Education*, vol. 2, pp. 68–70, 2011.
- [41] W. J. Huang, W. J. Qiu, and W. Q. Jiang, “A study of the relationships among the scores of Othello game and fourth graders’ reasoning ability, creativity ability, and problem-solving ability,” *Research and Development in Science Education Quarterly*, vol. 47, pp. 19–46, 2007.
- [42] R. L. Zheng, “Research on game design of elementary mathematics education for developing problem-solving ability,” *Journal of Jiamusi Education Institute*, vol. 8, pp. 1–51, 2014.
- [43] T. C. Halloluwa, D. Vyas, T. R. Sahama, P. Hewagamage, and H. Usoof, “Interaction design for tablet based edutainment systems for mathematical education of primary learner,” in *Proceedings of the 2015 Fifteenth International Conference on Advances in ICT for Emerging Regions (ICTer)*, Colombo, Sri Lanka, August 2015.
- [44] Y. H. Lee, N. Dunbar, K. Kornelson et al., “Digital game based learning for undergraduate calculus education: immersion, calculation, and conceptual understanding,” *International Journal of Gaming and Computer-Mediated Simulations*, vol. 8, no. 1, pp. 13–27, 2016.
- [45] H. T. Muñoz, R. F. Gesa, and S. Baldiris, “Augmented reality game-based learning for mathematics skills training in inclusive contexts,” *IE Comunicaciones: Revista Iberoamericana de Informática Educativa*, vol. 21, p. 4, 2015.
- [46] J. Vrugte, T. Jong, S. Vandercruyse, P. Wouters, H. Oostendorp, and J. Elen, “How competition and heterogeneous collaboration interact in prevocational game-based mathematics education,” *Computers & Education*, vol. 89, pp. 42–52, 2015.
- [47] H. S. Lee and J. W. Lee, “Mathematical education game based on augmented reality,” in *Technologies for E-Learning and Digital Entertainment*, pp. 442–450, Springer, Berlin, Germany, 2008.
- [48] V. E. Richard, “Digital game-based learning: it’s not just the digital natives who are restless,” *Educa*, vol. 41, pp. 16–30, 2006.
- [49] S. Tang, M. Hanneghan, and A. El Rhalibi, “Introduction to games-based learning,” in *Games-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices*, T. Connolly, M. Stansfield, and L. Boyle, Eds., IGI Global, Hershey, PA, USA, pp. 1–17, 2009.
- [50] M. Griffiths and M. N. O. Davies, “Research note Excessive online computer gaming: implications for education,” *Journal of Computer Assisted Learning*, vol. 18, no. 3, pp. 379–380, 2002.
- [51] B. Paras and J. Bizzocchi, “Game, motivation, and effective learning: an integrated model for educational game design,” in *Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in Play*, Vancouver, Canada, June 2005.

- [52] B. Kim, H. Park, and Y. Baek, "Not Just fun, but serious strategies: using meta-cognitive strategies in game-based learning," *Computers & Education*, vol. 52, no. 4, pp. 800–810, 2009.
- [53] K. A. Wilson, W. L. Bedwell, E. H. Lazzara et al., "Relationships between game attributes and learning outcomes: review and research proposals," *Simulation & Gaming*, vol. 40, no. 2, pp. 217–266, 2009.
- [54] J. M. Yien, C. M. Hung, G. J. Hwang, and Y. C. Lin, "A game based learning approach to improving students' learning achievements in a nutrition course," *Turkish Online Journal of Educational Technology*, vol. 10, no. 2, pp. 1–10, 2011.
- [55] L. J. Chu, C. L. Li, and J. L. Guo, "Influence of flow experience on learning effect in interactive multimedia materials," *Journal of Cultural Enterprise and Management*, vol. 6, pp. 1–24, 2011.
- [56] Y. H. Chen, Y. M. Yeh, K. C. Li, M. S. Li, and M. H. Qiu, "Compilation of primary school students' scientific attitude scale and related factors," *Journal of National Tainan Teachers College*, vol. 24, pp. 1–26, 1991.
- [57] Z. P. Cao and W. Z. Zhou, "Compilation of the national mathematics attitude scale," *In Proceedings of the Academic Symposium on Education in 1998-1999 Academic Year*, vol. 3, pp. 1211–1245, 1998.
- [58] J. Zhou and J. M. George, "When job dissatisfaction leads to creativity: encouraging the expression of voice," *Academy of Management Journal*, vol. 44, no. 4, pp. 682–696, 2001.
- [59] C. Y. Zhang and C. C. Dong, "To problem solving or not to problem solving? A comparison of different CAI effectiveness," *Chinese Journal of Science Education*, vol. 8, no. 4, pp. 357–377, 2000.
- [60] P. S. Georgilakis, G. A. Orfanos, and N. D. Hatziargyriou, "Computer-assisted interactive learning for teaching transmission pricing methodologies," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1972–1980, 2014.

Research Article

Sizing of the Propulsion System for a Heavy-Duty Fuel Cell Commercial Vehicle

Weiguang Zheng ^{1,2}, Weiwei Xin ^{2,4}, Enyong Xu ^{1,3}, Shuilong He ⁴, Jirong Qin,³
and Heng Wang⁴

¹State Key Laboratory of Digital Manufacturing Equipment and Technology, School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

²School of Mechanical and Automotive Engineering, Guangxi University of Science and Technology, Liuzhou 545616, China

³Commercial Vehicle Technology Center, Dong Feng Liuzhou Automobile Co., Ltd., Liuzhou 545005, China

⁴School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin, 541004, China

Correspondence should be addressed to Enyong Xu; xuey@dfzm.com

Received 1 October 2021; Revised 23 November 2021; Accepted 6 December 2021; Published 21 December 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Weiguang Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a methodology for the sizing of a heavy-duty fuel cell commercial vehicle. The parameters scanning model and the long-term stochastic drive cycle are adopted for this proposed sizing framework. The dynamic programming algorithm is employed as the energy management strategy to assess the performance of sizing. The efficacy of this framework is evaluated, and a detailed analysis for the hydrogen consumption is given in the results. In addition, a prediction analysis based on the price performance of the next decade is also given in this work.

1. Introduction

The energy and environment are two limitations for the transportation industry [1, 2]. The heavy-duty commercial vehicle is imperative to turn to new energy for its high fuel consumption and emission. As a fuel-renewable and environment-friendly on-board propulsion system, hydrogen fuel cells are of great significance for alleviating the current energy and environment dilemma.

Restricted by the dynamic properties of fuel cell, fuel cell vehicles (FCVs) are always equipped with at least two energy sources: a fuel cell stack and a battery pack or a supercapacitor, which are also known as fuel cell hybrid electric vehicles (FCHEVs) [3]. Therefore, FCHEV is generally treated as a particular electric vehicle which equipped with two or three power sources [4]. Many approaches are presented in the literature to optimize the vehicle operating cost, which can be divided into two levels: the optimization level and the design level. The optimal operation aims to find the optimal operating points for fuel cell to achieve the

minimum hydrogen cost, which is determined by the energy management strategy (EMS). The optimal design aims to determine the optimal sizes of vehicle components [5, 6].

The key role of EMS for a FCHEV is to decide the power splitting for energy sources at each instant, while at the same time, the drivability constraints of vehicle must be satisfied [7]. There have been many researches on EMS to improve the fuel economy. In general, rule-based EMS and optimization-based EMS are two approaches widely studied [8, 9]. Under current technology, rule-based EMS has been widely used in real vehicles' controller for its advantages in real time, simplicity, and cost. However, subjectivity and uncertainty of this strategy may lead the control rules far to the optimal rules [8]. On the other hand, thanks to the increasing computing power of on-board processors, more complex EMS has the potential to be applied in real vehicles; optimization-based EMS has become the focus of current academic research. Optimization-based EMS is often represented as optimization problems for constrained systems in finite-time domain; the common optimization algorithms

include dynamic programming (DP) [10], Pontryagin's minimum principle (PMP) [11], and genetic algorithm (GA) [12]. In addition, two optimization frameworks (equivalent consumption minimization strategy (ECMS) [13] and model prediction control (MPC) [14–16]) are also applied to this strategy to improve the control effect.

In addition to the EMS, many researchers focused their attention on the sizing of energy sources to improve vehicle performance. The main purpose of sizing of energy sources is to optimize the dynamic and fuel economy with fewer cost; reasonable parameters can maximize the potential of a vehicle's EMS [17]. With the development of electric vehicle, numbers of algorithms have been adopted to solve this problem. As early as 2004, Lukic and Emadi [18] verified the influence of drivetrain hybridization on energy consumption and dynamic performance. In their work, the hybridization factor was defined, and the EMS was implemented as charge sustaining. Kim and Peng [19] proposed a combined optimization method of EMS and parameters sizing for FCHEV. Different subsystem-scaling models were adopted to predict the characteristics of components on different sizes, and a near-optimal EMS was designed with the inspiration of stochastic dynamic programming results. In [20], the scholars constructed a stochastic drive cycle to simulate the randomness of real-world driving conditions, and ECMS was used as an online EMS; the equivalent factors of ECMS were adjusted with the proposed sizing methodology. In [21], a parameters' sizing methodology based on ordinal optimization and dynamic programming was proposed. Hu et al. [22] gave a compared analysis for two optimization-based EMS, and the influence of downsizing the battery was also studied. In their further study, the influence of driving pattern on sizing was studied in [23]; the comparison of three different electrochemical energy storage systems and the sizing framework for hybrid electric vehicle are also given [24, 25]. In addition, Karaođlan et al. [26] studied the influence of gear ratio on fuel economy and emissions.

For the EMS, optimization-based strategy has been the focus of current research; and for the parameters sizing, systematic algorithms are used to find the optimal/near-optimal sizes of components for FCHEV; rule-based EMS or optimization-based EMS are adopted as an associated problem for these algorithms. Although existing research has proposed a number of approaches for parameters sizing, these approaches are based on the few drive cycles or only one standard drive cycle, and parameters' sizing for hybrid heavy-duty commercial vehicles is little studied.

In this paper, a sizing methodology for a fuel-cell/battery commercial vehicle is presented. It is based on the DP-based EMS and parameters scanning model, and a long-term stochastic drive cycle is adopted for the sizing framework. The main contributions are as follows: (a) A long-term stochastic drive cycle that is based on the cluster and stochastic procedure is adopted, and the processes of generate stochastic drive cycle are also available. (b) A detailed analysis on oversizing the fuel cell stack and battery is given. (c) A predictive analysis is also given with considering the price performance in the next decade.

The remainder of this paper is organized as follows. In Section 2, the model of a heavy-duty fuel cell commercial vehicle is established. In Section 3, the design process of the long-term stochastic drive cycle is proposed. In Section 4, the DP-based EMS is introduced. In Section 5, the sizing methodology of propulsion system is presented. Section 6, provides results and discussion, and in Section 7, the conclusion of this paper is given.

2. Mathematical Model of an FCHEV

In general, for an FCHEV, fuel cell stack is employed as the main power source, and battery is adopted as an energy storage system, which is also known as the auxiliary source. The structure of selected FCHEV is shown in Figure 1. The electric motor is the only conversion equipment of electric energy and kinetic energy, fuel cell stack and battery pack composed of the propulsion system to provide power to permanent magnet synchronous motor and the motor's inverter (DC-AC), and a DC-DC converter is used to boost the voltage from fuel cell stack to DC bus. In this work, the powertrain system is divided into a combined propulsion subsystem and a transmission subsystem. The parameters sizing of propulsion system is the main work of this paper.

2.1. Vehicle and Drivetrain. The mathematical model building of FCHEV is reversed with the target of tracking drive cycle, and the impacts of lateral dynamics and rotating mass are ignored. At every moment of vehicle's operation, the torque T_w and speed ω_w of wheels are given by

$$\begin{aligned} T_w &= r \left(m \frac{dv}{dt} + \frac{1}{2} \rho C_d A_f v^2 + mgf \cos(\alpha) + mg \sin(\alpha) \right), \\ \omega_w &= \frac{v}{r}, \end{aligned} \quad (1)$$

where r is the rolling radius of the wheel, m the FCHEV's mass with full load; v the vehicle's velocity at the current moment; ρ the air density; A_f the frontal area; C_d the aerodynamic drag's coefficient; f the rolling resistance coefficient; α the road's inclination angle. The motor's torque T_m , speed ω_m , and power P_m are formulated as follows:

$$T_m = \begin{cases} \frac{T_w}{\eta_{fd} r_{fd}}, & T_w \geq 0, \\ \frac{T_w \eta_{fd}}{r_{fd}}, & T_w < 0, \end{cases} \quad (2)$$

$$\omega_m = \omega_w r_{fd},$$

$$P_m = \begin{cases} \frac{T_m \omega_m}{\eta_m}, & T_m \geq 0, \\ T_m \omega_m \eta_m, & T_m < 0, \end{cases}$$

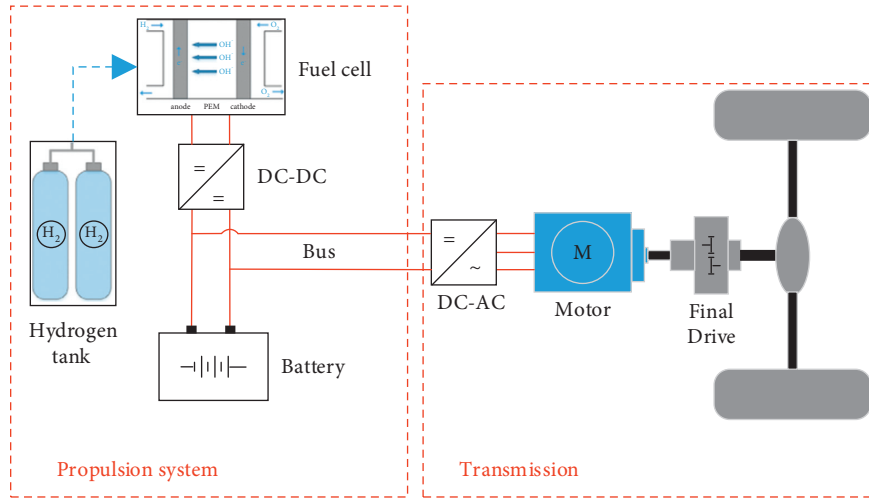


FIGURE 1: Block diagram of FCHEV. The blue dotted lines, red thin solid lines, and black thick solid lines represent the hydrogen gas flow channel, the electrical connections and mechanical connections of FCHEV, respectively.

where η_{fd} and r_{fd} are the efficiency and gear ratio of the final drive, respectively.

For this work, the motor power is defined as positive during traction while negative during braking. The propulsion system consists of fuel cell stack and battery pack outputs the power demand of the motor, as shown in the following formula:

$$P_m = P_{\text{stack}} + P_{\text{pack}}, \quad (3)$$

where P_{stack} and P_{pack} are the power of fuel cell stack and battery pack. Parameters of the heavy-duty fuel cell commercial vehicle are listed in Table 1.

2.2. Fuel Cell Stack. Fuel cell is an electrical device that converts chemical energy into electricity. Proton exchange membrane fuel cell (PEMFC) is the most common on-board propulsion system for fuel cell vehicles. A complete on-board PEMFC system includes fuel cell stack, hydrogen storage system, hydrogen circuit, air circuit, water circuit, and coolant circuit [3]. The models for these systems are very complicated. This work is mainly concerned with the power-split of the bus but not the detailed conversion process of fuel cells. Therefore, a simple efficiency graph model is used for the simulation.

In this selected model, hydrogen consumption and efficiency of the fuel cell stack are both formulated as the function of power demand; these datasets are obtained through testing, which have been provided in ADVISOR (FC_ANL50H2) [27]. Scaling models are adopted to facilitate parameters' sizing, as shown in Figure 2.

2.3. Battery Pack. Benefiting from the high power density and high reliability, lithium-ion batteries have become the most commonly used energy storage system for FCHEV. The purpose of adding a battery pack into the propulsion system is to improve the drivability and efficiency performance of the vehicle.

TABLE 1: Parameters of the FCHEV.

Parameters	Value
Vehicle total mass (kg)	9000
Air density (kg/m^3)	1.2
Aerodynamic drag coefficient	0.563
Vehicle frontal area (m^2)	6.6454
Wheel radius (m)	0.413
Gear ratio	6.67
Efficiency of transmission (%)	98
Rolling resistance coefficient	0.015
Efficiency of converter (%)	98
Efficiency of inverter (%)	98
Efficiency of charge (%)	90
Efficiency of discharge (%)	100

For the model of battery pack, the circuit model is used for simulation. In this model, the battery is equivalent to a circuit with an ideal voltage source and an internal resistance. When the battery pack is operating, its current and power are formulated as follows:

$$I_{\text{bat}} = \frac{V_{oc} - \sqrt{V_{oc}^2 - 4R_{\text{int}}P_{\text{pack}}}}{2R_{\text{int}}}, \quad (4)$$

where V_{oc} is the open circuit voltage, I_{bat} the current, R_{int} the internal resistance, and P_{pack} the power of battery pack. These parameters can be obtained from equations (5) to (7).

$$P_{\text{pack}} = P_{\text{bat}}\eta_{\text{bat}}, \quad (5)$$

$$V_{oc} = \frac{V_{oc}n_s}{n_p}, \quad (6)$$

$$R_{\text{int}} = \frac{r_{\text{int}}n_s}{n_p}, \quad (7)$$

where η_{bat} is the conversion efficiency for charge/discharge, n_s the series number of cells, and n_p the parallel number of

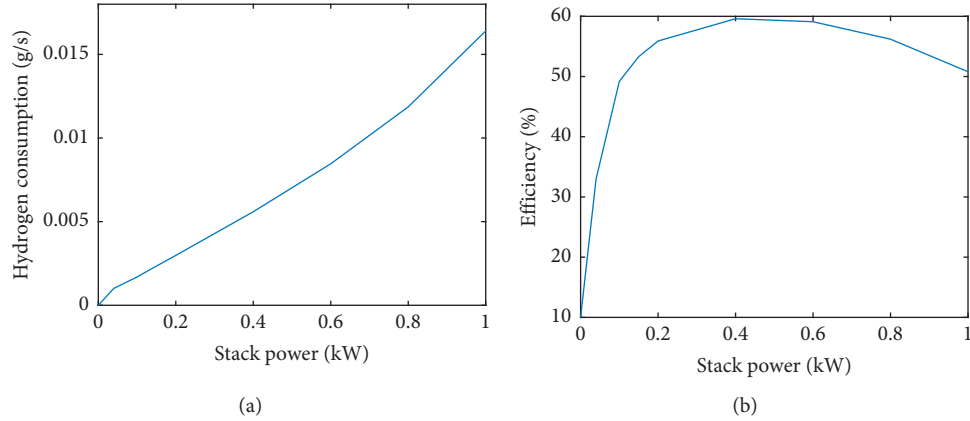


FIGURE 2: Scaling models of fuel cell stack (on the basis of a fuel cell stack with 50kW peak power on ADVISOR). (a) Hydrogen consumption at different stack power. (b) Efficiency at different stack power.

cells. The state of charge (SOC) of battery at time step k is defined as

$$\text{soc}(k) = \text{soc}(k-1) - \frac{I_{\text{bat}}}{Q_{\text{bat}}}, \quad (8)$$

where Q_{bat} is the capacity of battery pack.

R_{int} and P_{pack} are the functions of SOC, as shown in Figure 3.

It is worth mentioning that the battery pack is composed of strings connected in parallel, with each string containing the same number of cells in series. More detailed descriptions about configuration and modelling of battery pack are available in [23].

2.4. Drive Motor. For an FCHEV, the drive motor's function mainly includes the following two parts: (a) converting the electrical energy from bus to mechanical energy to drive the vehicle and (b) recycling the vehicle's braking power to charge the battery. A permanent magnet synchronous motor (PMSM) is adopted for its compact and high efficiency. The model of drive motor is based on its efficiency map, as shown in Figure 4.

3. Long-Term Stochastic Drive Cycle

The drive cycle refers to the driving characteristics of a certain type of vehicle (passenger car, bus, and cargo vehicle, etc.) in a specific condition (urban, suburban, highway, etc.); it is generally a set of points representing speed versus time. There have been many systems of typical driving cycles around the world.

Although the representativeness of cycles will be considered in the design process of typical drive cycle, the optimization results obtained under few drive cycles cannot satisfy the randomness of real driving conditions [20]. Therefore, a long-term stochastic drive cycle is built to overcome this drawback. The constructed drive cycle is based on the typical drive cycles of heavy commercial vehicles, and cluster analysis and random process are adopted. The flowchart of this procedure is shown in Figure 5.

In this work, eight typical drive cycles for commercial vehicles are introduced as the DC library, which are numbered and shown in Figure 6. Twelve characteristic parameters are used for clustering, and the corresponding descriptions of these parameters are listed in Table 2. The characteristic parameters for each drive cycle are calculated and listed in Table 3. These parameters will be the raw data of clustering.

Hierarchical clustering and Euclidean distance are adopted to obtain the similar degree of the different drive cycles. The clustering results are shown in Figure 7. The library is divided into 5 classes with the benchmark of similarity to 1: the first category includes drive cycles 1, 4, 6, and 8, and drive cycles 2, 7, 3, and 5 are classified as the second to fifth categories, respectively. The drive cycles 8 (CHTC_HT), 2 (WVUCITY), 7 (NYCOMP), 3 (WVUINTER), and 5 (UDDSSHCV) are selected as the representatives of each category. These five drive cycles will be regarded as the originals in the long-term stochastic drive cycle.

The next step for this work is to build the combined drive cycle, which is generated from the original drive cycles. After that, the combined drive cycle will be divided into microtrips and microidles. A microtrip is a segment of the drive cycle where the velocity is not equal to zero, and a microidle is the segment where the velocity is equal to zero. The relationship between microtrips and microidles is depicted in Figure 8. Then, the divided segments are added to the stochastic drive cycle randomly. In this process, the microtrips and microidles are scaled with a stochastic procedure. The scaling processes for microtrips are for their velocity, and for microidles, it is for their length of time, the scaling factors are both defined as a stochastic number from 0.8 to 1.2. To follow the objective laws for drive cycles, the microtrip and microidle should be added in turn, which means after adding a microtrip, a microidle must be added as its follow-up. In addition, the long-term stochastic drive cycle can more effectively reflect the randomness of the drive cycle, the total length of stochastic drive cycle is set as 10 times that in the combined drive cycle in this work.

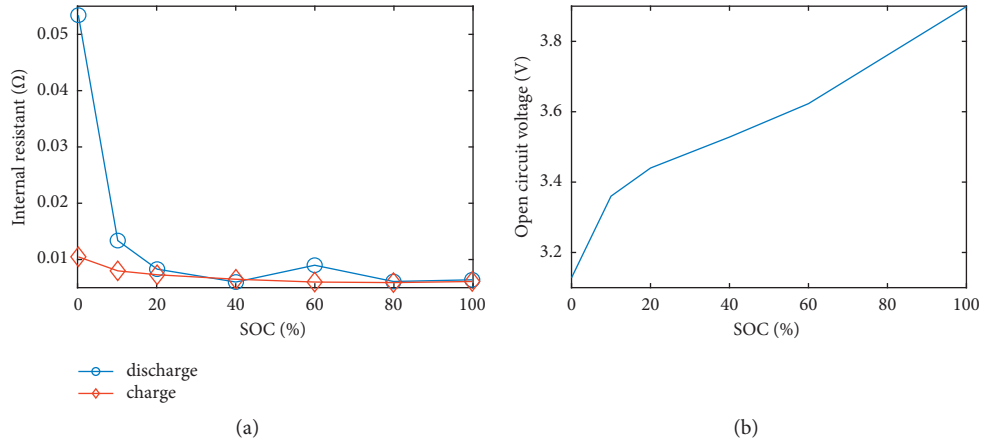


FIGURE 3: Circuit voltage and internal resistance of the cell. (a) Resistances of cell for different SOC in charge and discharge state. (b) Open circuit voltage of cell for different SOC.

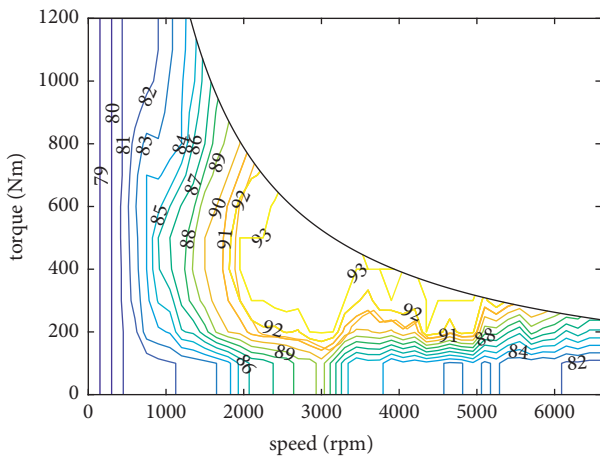


FIGURE 4: Efficiency map of PMSM.

Another noteworthy point for this process is that after the scaling of a microtrip, three indexes should be detected: the maximum speed, maximum acceleration, and maximum deceleration. If any index is higher than the combined drive cycle, the scaling process should be repeated until satisfied.

After completing the construction of long-term stochastic drive cycle, evaluation indexes are introduced to evaluate the rationality. Figure 9 illustrates the distribution of two drive cycles; Table 4 depicts the comparison of characteristics of combined drive cycle and long-term stochastic drive cycle. The above charts show that the long-term stochastic drive can accurately reflect the characteristics of original cycles.

4. DP-Based EMS

In general, the EMS for FCHEV has a great impact on efficiency. In this work, an optimal EMS based on dynamic programming algorithm is adopted to assess the economic potential of sizing.

Dynamic programming is a numerical method for solving multistage decision-making problems. It can provide

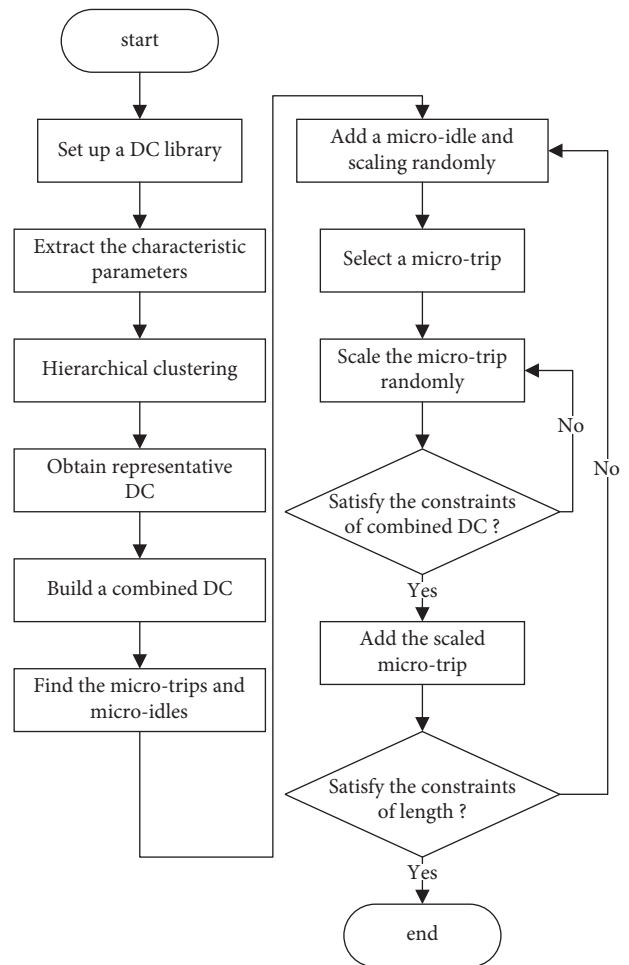


FIGURE 5: Flowchart of long-term stochastic drive cycle generation. DC: drive cycle.

the optimal solution for problems of any complexity level in the limits of computational capabilities [7].

For a controlled multistage decision-making problem, the state transfer function can be described as

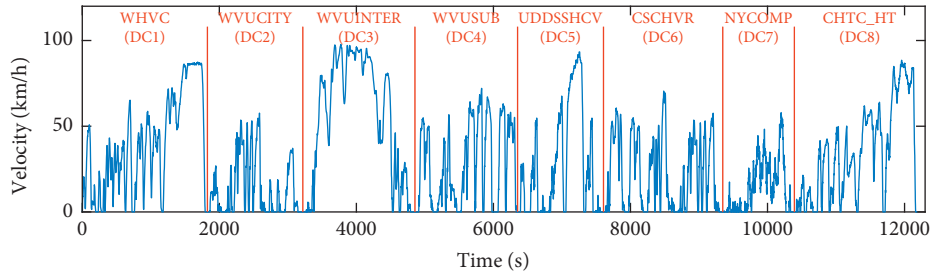


FIGURE 6: The serial number of the selected drive cycles.

TABLE 2: The characteristic parameters used for clustering.

Parameter	Denotation
Average velocity (m/s)	\bar{v}
Average acceleration (m/s ²)	\bar{a}
Average deceleration (m/s ²)	\bar{d}
Standard deviation of velocity	σ_v
Maximum velocity (m/s)	v_{\max}
Maximum acceleration (m/s ²)	a_{\max}
Maximum deceleration (m/s ²)	d_{\max}
The percentage of time in speed interval 0–20 km/h (%)	θ_{0-20}
The percentage of time in speed interval 20–40 km/h (%)	θ_{20-40}
The percentage of time in speed interval 40–60 km/h (%)	θ_{40-60}
The percentage of time in speed >60 km/h (%)	$\theta_{60-\max}$
The percentage of time in idle state (%)	θ_0

TABLE 3: The characteristic parameters that are calculated for each drive cycle.

Denotation	DC1	DC2	DC3	DC4	DC5	DC6	DC7	DC8
\bar{v}	11.25	3.78	15.22	7.19	8.42	6.07	3.92	9.62
\bar{a}	0.23	0.30	0.20	0.33	0.48	0.39	0.47	0.22
\bar{d}	0.28	0.39	0.21	0.42	0.58	0.46	0.54	0.29
σ_v	7.81	4.58	9.77	6.56	8.87	5.82	4.23	7.61
v_{\max}	24.25	16.01	27.15	20.03	25.93	19.58	16.09	24.58
a_{\max}	0.82	1.14	1.42	1.29	1.96	1.16	2.06	1.22
d_{\max}	1.17	2.24	1.86	2.16	2.07	1.79	1.95	1.25
θ_{0-20}	21.93	41.26	15.85	24.86	16.49	31.56	33.98	22.56
θ_{20-40}	23.10	17.19	9.21	13.51	15.83	19.43	25.63	22.67
θ_{40-60}	21.65	11.29	11.83	28.53	12.06	24.76	7.28	20.61
$\theta_{60-\max}$	26.49	0	53.84	7.93	22.34	2.64	0	19.89
θ_0	6.38	30.26	9.27	25.17	33.27	21.62	33.11	14.28

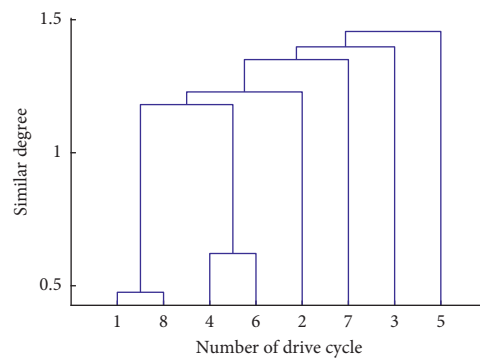


FIGURE 7: Results of hierarchical clustering.

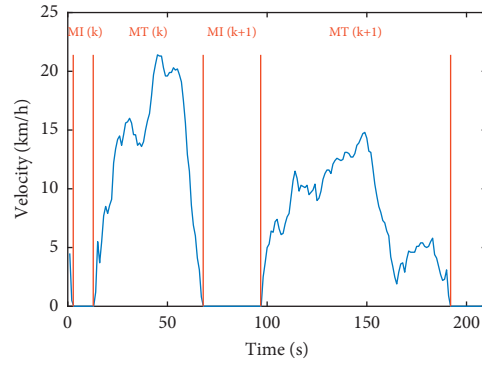


FIGURE 8: Relationship of the microtrips and microidles. The MI(k) and MT(k) represent the kth microidle and microtrip, respectively.

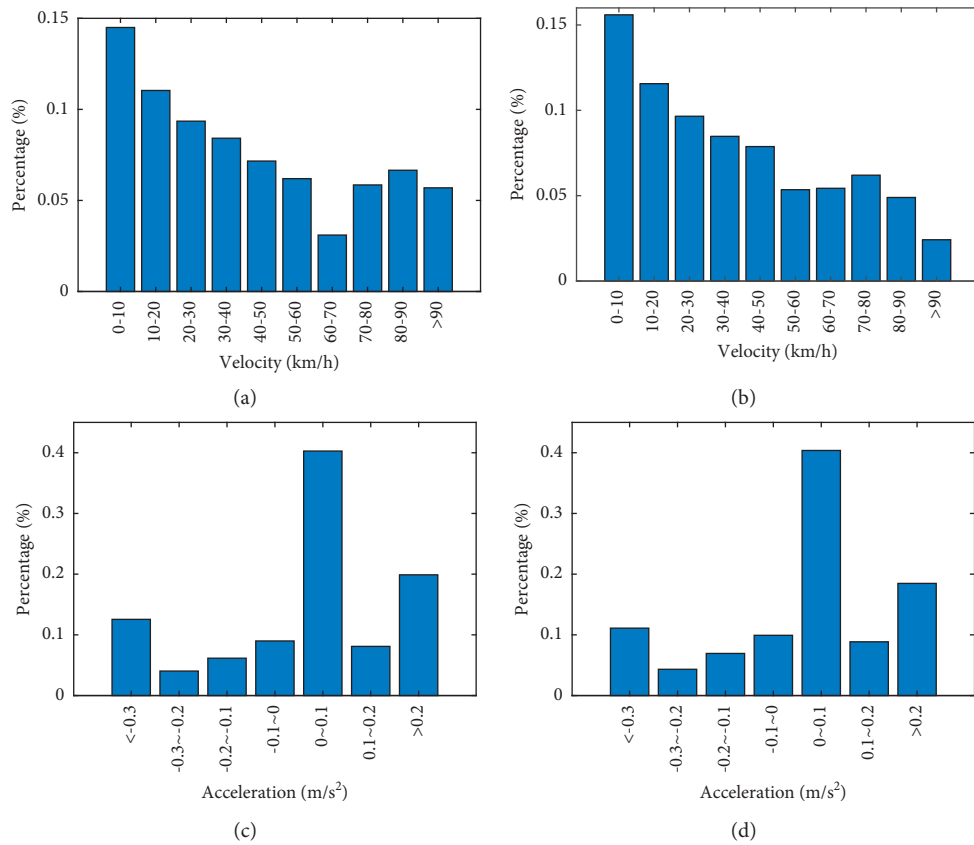


FIGURE 9: The distribution of velocity and acceleration for two drive cycles. (a) The percentage of speed for combined drive cycle. (b) The percentage of speed for long-term stochastic drive cycle. (c) The percentage of acceleration for combined drive cycle. (d) The percentage of acceleration for long-term stochastic drive cycle.

TABLE 4: The characteristics of two drive cycles.

Characteristics	Combined DC	Stochastic DC
Duration (s)	6939	69393
Idle time (%)	22.03	24.38
Maximum velocity (m/s)	27.15	27.08
Average velocity (m/s)	8.73	6.85
Standard deviation of velocity	8.66	7.06
Maximum acceleration (m/s ²)	2.07	2.06
Maximum deceleration (m/s ²)	3.24	3.17
Average acceleration (m/s ²)	0.29	0.28
Average deceleration (m/s ²)	0.35	0.34

$$s(k+1) = f_k(s(k), u(k)), \quad (9)$$

where $s(k)$ is the state variable at time k and $u(k)$ is the control variable. For a control policy $u = \{u_0, u_1, \dots, u_{N-1}\}$, the cost in time step N is defined as

$$J(s_0, u) = L_N(s_N) + \sum_{k=1}^{N-1} L_k(s_k, u_k), \quad (10)$$

where L_k is the instantaneous cost function. The optimal cost is

$$J^*(s_0) = \min(J(s_0, u), u). \quad (11)$$

The optimal control policy $u^* = \{u_1^*, u_2^*, \dots, u_{N-1}^*\}$.

For the DP-based EMS for FCHEV, the SOC of battery is selected as the state variable, and the output power of fuel cell stack is adopted as the control variable. The state transfer function is defined as follows:

$$\text{SOC}(k+1) = \text{SOC}(k) + h_k(P_{\text{stack}}), \quad (12)$$

where $h_k(P_{\text{stack}})$ is the change rate of state in time k , which can be obtained with equations (7) to (11).

For the structure of non-plug-in FCHEV, a constraint of terminal SOC is required to sustain the energy storage system. In principle, the constraint of terminal SOC can be taken into account in two different ways, namely, as a soft constraint or a hard constraint. In this study, a hard constraint is adopted by requiring that the energy storage stored at the end of the mission equal to the value at the start of the mission, as shown in equation (13). A more detailed explanation will be found in [28–30]. Other parameters under constraints are listed in equation (14). In particular, the SOC_{\min} and SOC_{\max} are set to 0.4 and 0.8, respectively.

$$\text{SOC}(\text{end}) = \text{SOC}(0), \quad (13)$$

$$\begin{cases} T_{m_min} \leq T_m \leq T_{m_max}, \\ \omega_{m_min} \leq \omega_m \leq \omega_{m_max}, \\ I_{\text{bat_min}} \leq I_{\text{bat}} \leq I_{\text{bat_max}}, \\ P_{\text{bat_min}} \leq P_{\text{bat}} \leq P_{\text{bat_max}}, \\ P_{\text{stack_min}} \leq P_{\text{stack}} \leq P_{\text{stack_max}}, \\ \text{SOC}_{\min} \leq \text{SOC} \leq \text{SOC}_{\max}. \end{cases} \quad (14)$$

5. Structure of Parameters' Sizing

The objective of sizing is to obtain the near-optimal sizes of maximum power of fuel cell stack ($P_{\text{stack_max}}$) and the number of battery cell (n_{bat}). The sizes of propulsion should satisfy the following objectives: (a) meet the drivability and (b) minimize the operating cost of the vehicle during its life cycle.

For drivability, passenger cars typically have higher velocity and acceleration targets, while for commercial vehicles, equipped with large load mass, the vehicle's climbing performance always receives more attention. Therefore, the sizing method is tested with three subobjects: (a) The peak power of fuel cell stack alone must be able to sustain the

maximum velocity; (b) the peak power with fuel cell stack and battery pack together must meet the maximum climbing slope index at a given velocity; (c) the power with fuel cell stack and battery pack together should fulfil the power demand of long-term stochastic drive cycle. The first subobject is used to obtain the low limiting value of $P_{\text{stack_max}}$, and the others are used to determine n_{bat} . It should be noted that for the subobject (c), the verification results are related to the selected EMS, DP-based strategy is employed for this work, which has been introduced in Section 4.

This sizing methodology is based on the generated long-term stochastic drive cycle, a parameters' scanning model is added for sizing, and the dynamic programming algorithm is adopted for the EMS. The flowchart of this proposed methodology is shown in Figure 10.

In this sizing methodology, the range of fuel cell stack's peak power is divided into i_{\max} independent nodes; the range of battery cell's number is divided into j_{\max} independent nodes; these two groups of nodes constitute a crossover model with $(i_{\max} \times j_{\max})$ crossover points, and for each crossover point, there is an FCHEV model corresponding to it. The optimal cost for each size is obtained by DP-based EMS. In addition, the low limiting value of $P_{\text{stack_max}}$ should meet the power demand of the vehicle at a maximum velocity (equation (15)), and the peak-powers provided by the battery and fuel cell should meet the requirements of the vehicle's climbing performance (equation (16)). The peak power of the battery pack can be expressed as the product of its capacity and the maximum discharge rate equation (17):

$$P_{\text{stack_max}} \eta_{\text{con}} \geq \frac{P_v}{\bar{\eta}_m \eta_{\text{inv}}}, \quad (15)$$

$$P_{\text{stack_max}} \eta_{\text{con}} + P_{\text{bat_max}} \eta_{\text{bat}} \geq \frac{P_c}{\bar{\eta}_m \eta_{\text{inv}}}, \quad (16)$$

$$P_{\text{bat_max}} = Q_{\text{bat}} C_{\text{max}}, \quad (17)$$

where P_v is the power demand for maximum velocity, P_c is the power demand for climbing, $\bar{\eta}_m$ is the average efficiency of motor, η_{bat} is the charge/discharge efficiency of battery, and η_{inv} and η_{con} are the efficiency of DC-AC inverter and DC-DC converter, respectively.

6. Results and Discussion

This simulation is carried out with MATLAB; the interval is set to 10 kW and 25 for the fuel cell stack's peak power and the battery cells' number, respectively.

Under the current technical, lithium-ion batteries can maintain a discharge time of 5 s at the peak discharge rate of 30 C [31]. However, high rate discharge will seriously affect the life of the battery. In this paper, the maximum discharge rate of the battery is set as 20 C. In addition, the capacity of the battery cell used in the simulation is 3 Ah. Therefore, on the basis of equations (18) to (20), the minimum constraints of $P_{\text{stack_max}}$ and n_{bat} will be estimated, as shown in Figure 11.

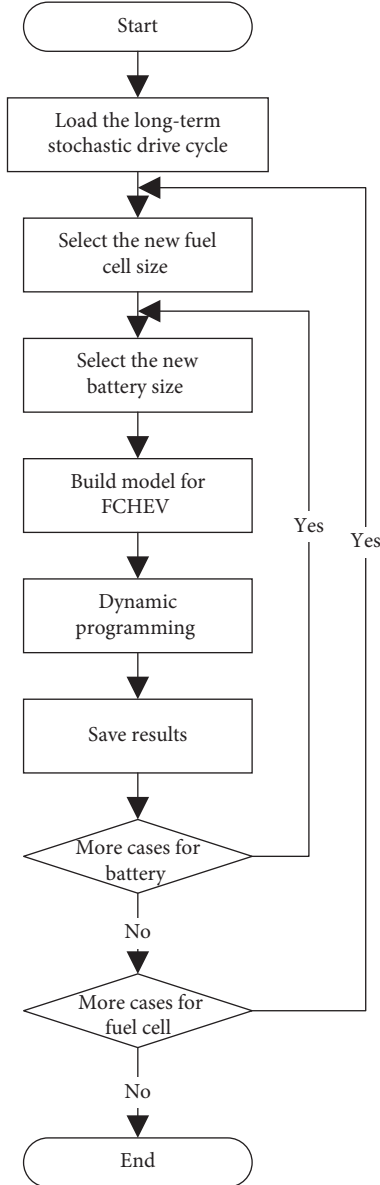
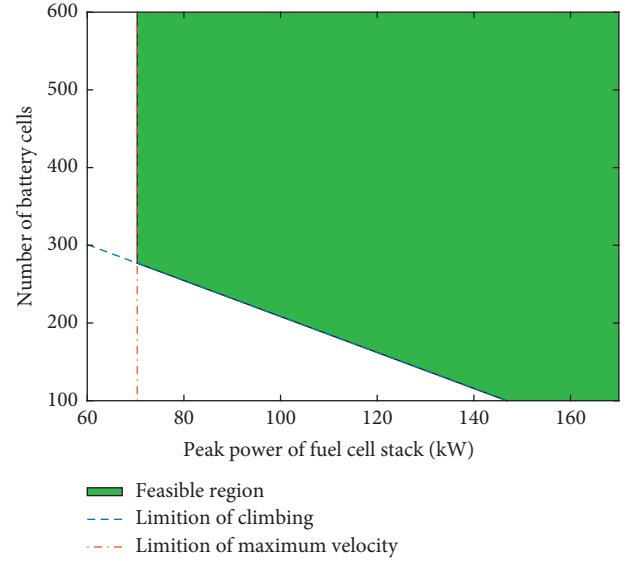


FIGURE 10: Flowchart of the sizing methodology.

Figure 12 shows the hydrogen consumption of vehicles with different sizes under long-term stochastic drive cycle. For a fair comparison, all the hydrogen costs are computed with the SOC deviation. Two conclusions can be obtained according to this figure:

- (1) Starting from $P_{\text{stack_max}} = 70$ kW, the hydrogen consumption decreases with the increasing of the peak power of fuel cell stack, but it increases when the peak power exceeds 120 kW
- (2) For a determined $P_{\text{stack_max}}$, the hydrogen consumption decreases with the increasing of the number of battery cells, but the slope of decline gradually decreases

To understand the causes of the changing of hydrogen consumption, we consider three relevant performance


 FIGURE 11: The minimum constraints of the $P_{\text{stack_max}}$ and n_{bat} .

indicators: the global propulsion efficiency, the braking-recovery-energy to negative-load-power ratio, and the global fuel cell stack efficiency.

The global propulsion efficiency is formulated as follows:

$$\eta_{g,\text{pro}} = \frac{\sum_{i=0}^N P_{\text{veh}}^+(i)\Delta t}{LHV_{H_2} \sum_{i=0}^N m_{H_2}(i)\Delta t} \times 100, \quad (18)$$

where P_{veh}^+ is the positive vehicle power, N is the length of stochastic drive cycle, m_{H_2} is the hydrogen consumption, and Δt is the simulation step.

The braking-recovery-energy to negative-load-power ratio is defined as follows:

$$\gamma_{re} = \frac{\sum_{i=0}^N P_{re}(i)\Delta t}{\sum_{i=0}^N P_{\text{veh}}^-(k)\Delta t} \times 100, \quad (19)$$

where P_{veh}^- is the negative vehicle power and P_{re} is the braking energy recovered by battery pack, which is formulated as follows:

$$P_{re} = \begin{cases} P_{\text{bat}}, & \text{if } P_{\text{veh}} < 0, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

The global fuel cell stack efficiency is defined as follows:

$$\eta_{g,\text{FC}} = \frac{\sum_{i=0}^N P_{\text{stack}}(i)\Delta t}{LHV_{H_2} \sum_{i=0}^N m_{H_2}(i)\Delta t} \times 100. \quad (21)$$

To make the figure clearer, four types of fuel cell stack peak power are selected for this work: $P_{\text{stack_max}} = 70$ kW, $P_{\text{stack_max}} = 100$ kW, $P_{\text{stack_max}} = 120$ kW, and $P_{\text{stack_max}} = 170$ kW. The rest results of other types are visible in Figure 13.

Figure 14 shows the effect of $P_{\text{stack_max}}$ and n_{bat} on defined relevant performance indicators. Figure 14(a) shows the maximum global propulsion efficiency is approximately 55.56%, which is reached when $P_{\text{stack_max}} = 120$ kW and $n_{\text{bat}} > 1000$. Increasing or decreasing the type of fuel cell will

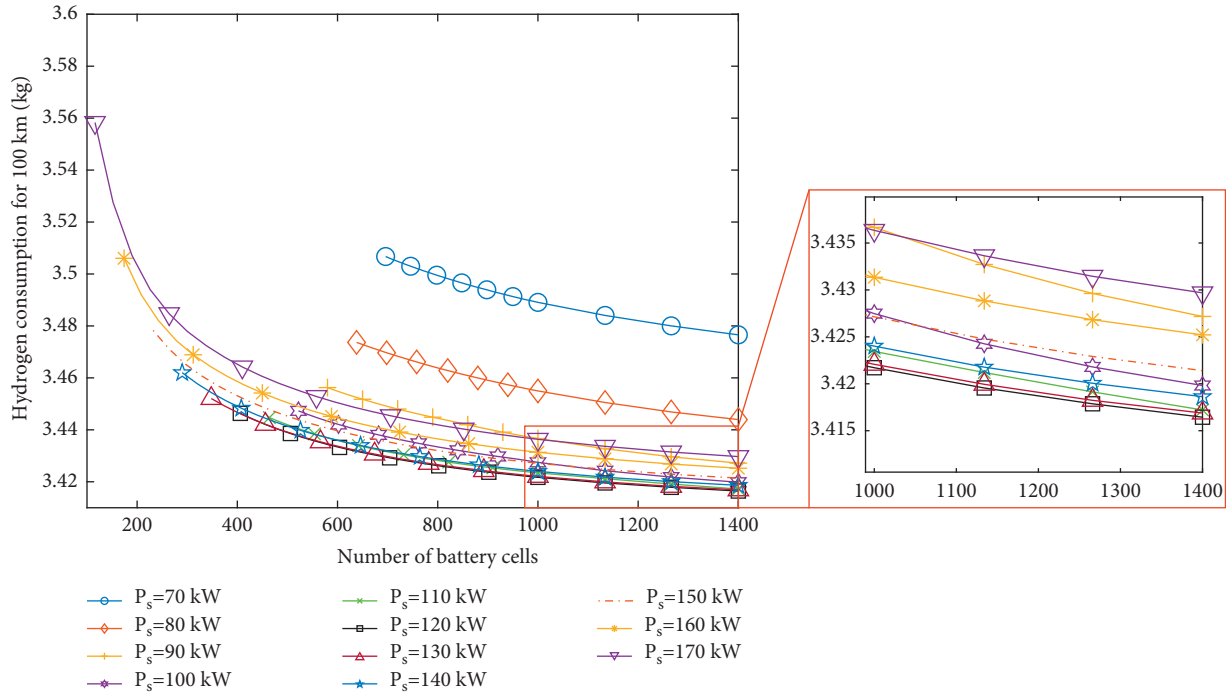


FIGURE 12: Hydrogen consumption in different sizes (P_s ; P_{stack_max}).

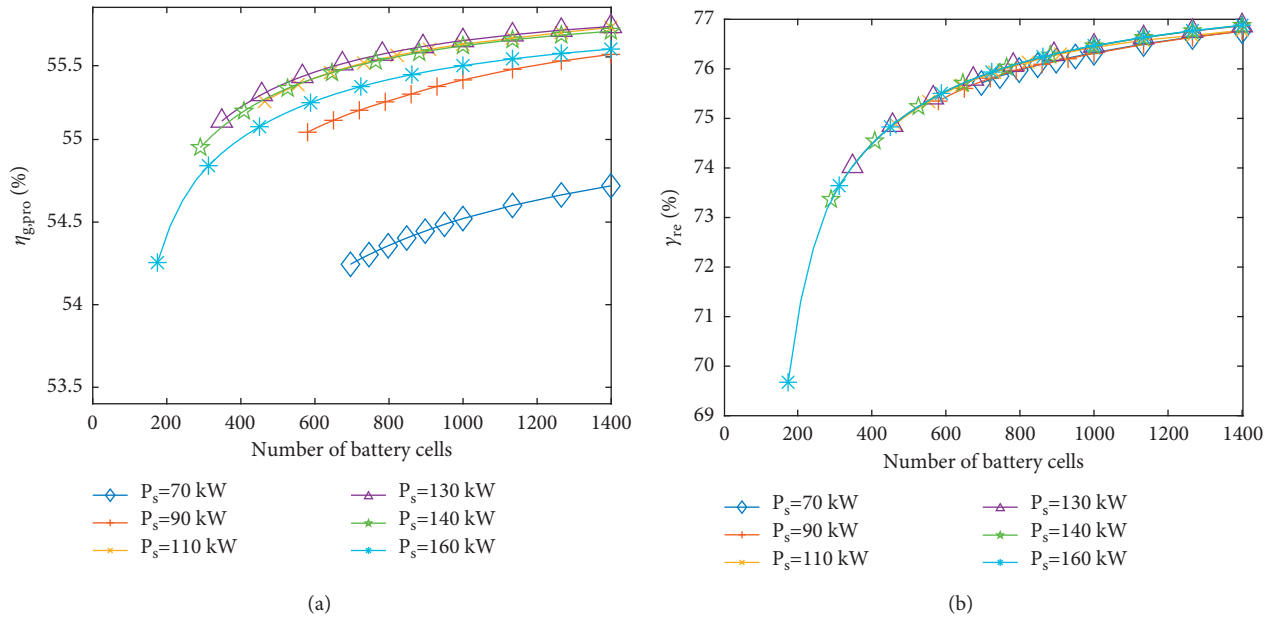


FIGURE 13: Continued.

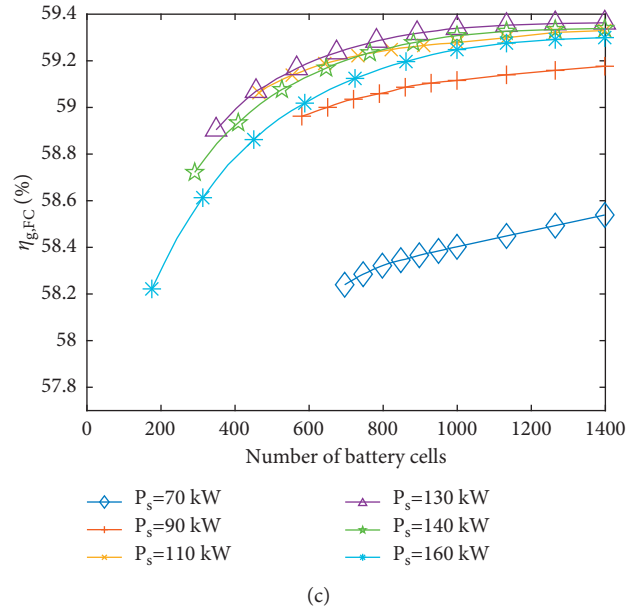


FIGURE 13: Comparative results of optimal costs.

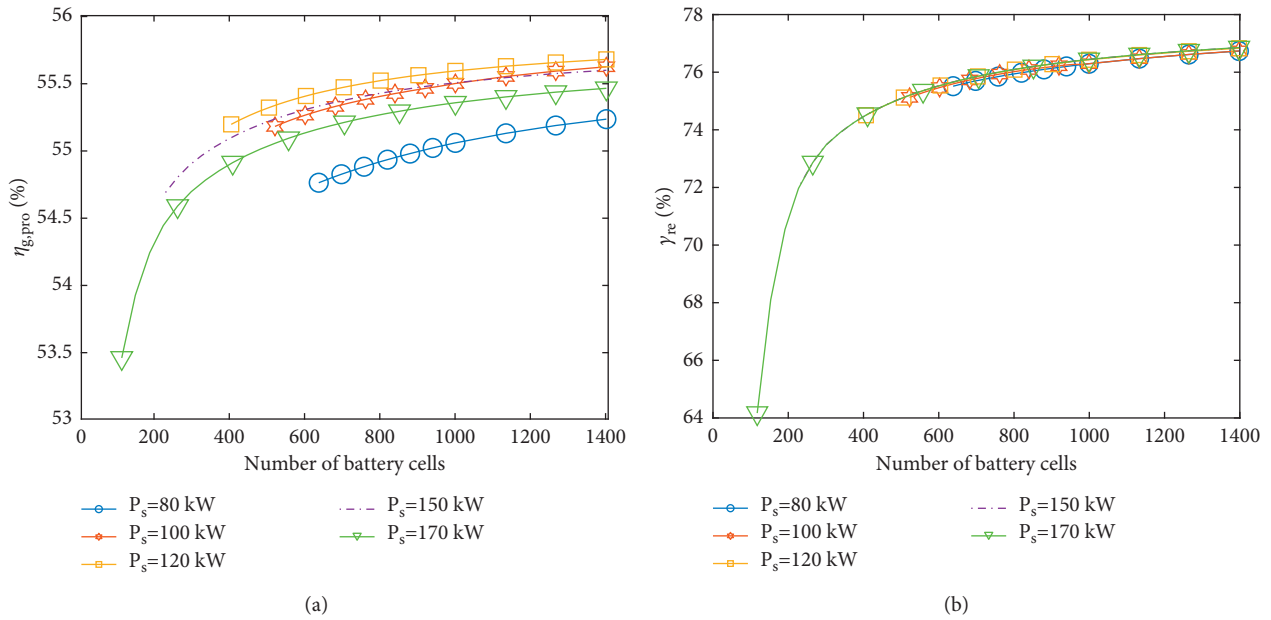


FIGURE 14: Continued.

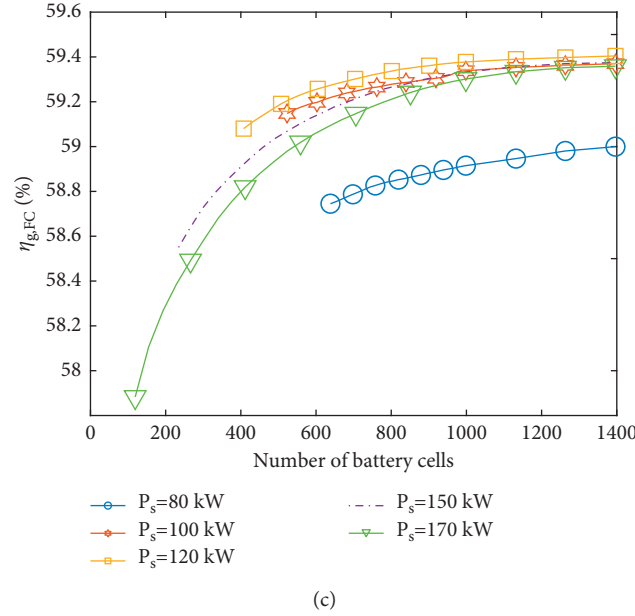


FIGURE 14: The test results of performance indicators with increasing cell number in different fuel cell sizes. (a) The global propulsion efficiency. (b) The braking-recovery-energy to negative-load-power ratio. (c) The global fuel cell stack efficiency.

worsen this indicator. On the other hand, increasing the number of battery cells can improve the global efficiency, but its growth trend gradually slows down and eventually to level. Figure 14(b) shows the size of fuel cell stack has less affected in braking-recovery-energy, while the number of battery cells has a great impact on this indicator: when $n_{\text{bat}} < 400$, limited by the physical constraints (e.g., the maximum SOC or maximum voltage), the battery pack is not enough to recover all the braking energy. This means that friction brakes will dissipate power from fuel cell stack, which is not friendly to the global efficiency. Figure 14(c) shows that the global fuel cell stack efficiency performs best when $P_{\text{stack_max}} = 120$ kW, and the increase in n_{bat} will benefit this indicator. Similarly, the growth trend gradually slows down with the increase of n_{bat} .

The results in Figure 10 clarifies that the reduction of hydrogen can be achieved with an appropriate $P_{\text{stack_max}}$, and the increase of n_{bat} will also benefits the fuel economy. It is beneficial to reduce the running cost, but the sharp increase in vehicle's price led by the increase of fuel cell and battery is unacceptable for both manufactures and users. Vehicle's price should also be considered as a further indicator.

To balance the cost of hydrogen consumption and component costs, a cost function which is defined as the cost per 100 km of the propulsion system is proposed, as shown in equation (18).

$$J^{100} = J_{\text{stack}}^{100} + J_{\text{bat}}^{100} + J_{H_2}^{100}. \quad (22)$$

In this function, the vehicle's cost J^{100} is divided into fuel cell stack cost J_{stack}^{100} , battery cost J_{bat}^{100} , and hydrogen cost $J_{H_2}^{100}$. The fuel cell stack cost and the battery cost are cycle-normalized cost, as expressed in equations (19) and (20) [28, 29]. In addition, the cost of hydrogen is the value of hydrogen

consumed when the vehicle's total driving distance reaches 100 km in the constructed long-term stochastic drive cycle, as shown in equation (21).

$$J_{\text{stack}}^{100} = \frac{d}{s y_v} \left(1 + P_c \frac{y_v + 1}{2} \right) j_{\text{stack}} P_{\text{stack_max}},$$

$$J_{\text{bat}}^{100} = \frac{d}{s y_v} \left(1 + P_c \frac{y_v + 1}{2} \right) j_{\text{bat}} Q_{\text{bat}}, \quad (23)$$

$$J_{H_2}^{100} = \frac{d}{X_{dc}} j_{H_2} m_{H_2},$$

where d is the length of cycle, X_{dc} is the distance under a stochastic drive cycle, s is the average travelled distance of the vehicle per year, y_v is the vehicle lifetime, and P_c is the yearly interest rate. Here, it is assumed that there are no components replaced during three-year operation, and the corresponding travelled distance is fifty thousand kilometres per year. These parameters above are given in Table 5. j_{stack} is the fuel cell stack price per kilowatt, j_{bat} is the battery price per kilowatt hour, and j_{H_2} is the hydrogen price per gram. Three price conditions on the years 2020, 2025, and 2030 [32] are chosen for this study, as shown in Table 6.

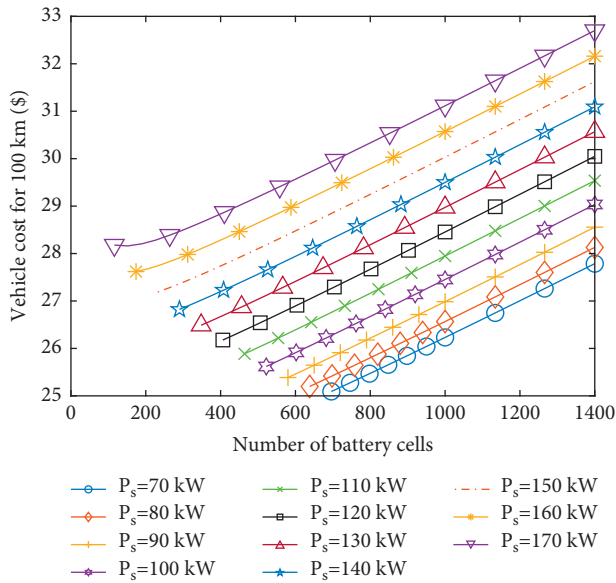
Figure 15 shows the price impact on vehicle's total cost. Figure 14(a) shows the vehicle total cost with the price in 2020. In this case, the unit prices of fuel cells and battery cell are higher; the optimal cost is 25.07 US dollars per 100 km with the corresponding size to $P_{\text{stack_max}} = 70$ kW and $n_{\text{bat}} = 696$. The cost previews of 2025 and 2030 are shown in Figure 14(b) and 14(c) respectively. On the preview of 2025, benefiting from the production increases and technology improves, the unit price of fuel cell and battery will decrease rapidly, and the optimal cost will drop to 21.46 US dollars per 100 km with the corresponding size to $P_{\text{stack_max}}$

TABLE 5: The parameters of simulation.

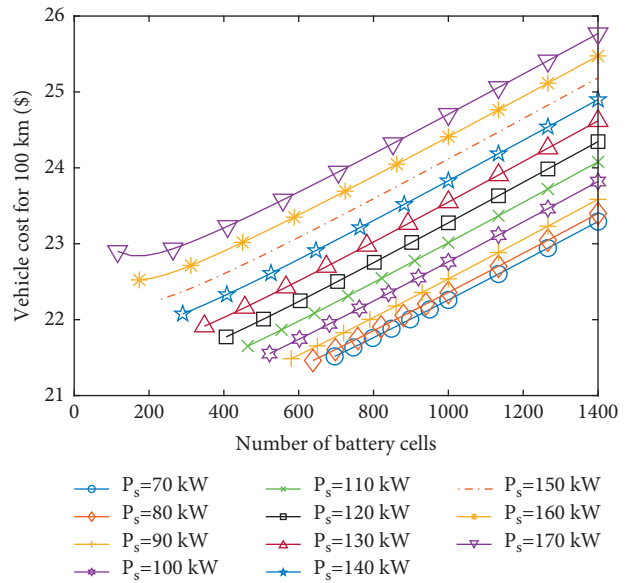
Description	Parameters	Value
Length of cycle (km)	d	100
Yearly travelled distance (km)	s	5e4
Vehicle lifetime (years)	y_v	3
Yearly interest rate (%)	P_c	5
Stochastic cycle distance (km)	X_{dc}	269

TABLE 6: The price trends of components cost.

Year	Fuel cell price (\$/kW)	Battery price (\$/kW)	Hydrogen price (\$/kg)
2020	70.41	256	5.32
2025	37.00	173	5.05
2030	29.84	117	5.03



(a)



(b)

FIGURE 15: Continued.

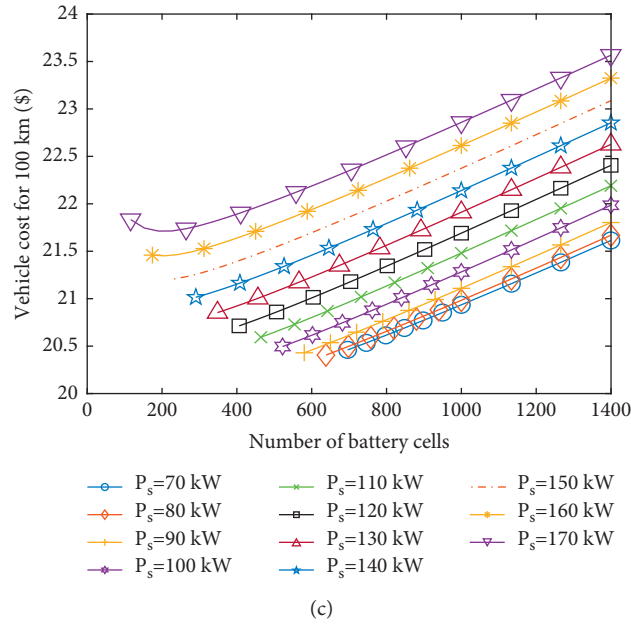


FIGURE 15: Performance indicators with increasing cell number in different fuel cell sizes. (a) The global propulsion efficiency. (b) The braking-recovery-energy to negative-load-power ratio. (c) The global fuel cell stack efficiency.

TABLE 7: The optimal sizes under different price conditions.

	2020	2025	2030
Optimal cost (\$/100 km)	25.07	21.46	20.41
Fuel cell size (kW)	70	80	80
Cell number	696	640	640
String number	4	4	4
Bus nominal voltage	626	576	576
Cell number per string	174	160	160

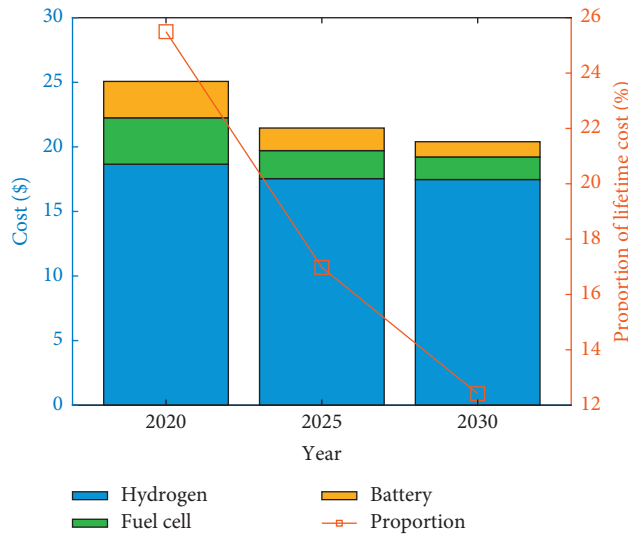


FIGURE 16: Vehicle costs for 100 km in different price conditions. (a) In 2020. (b) In 2025. (c) In 2030.

= 80 kW and $n_{bat} = 640$. In 2030, with the further decrease of component prices, the optimal cost will drop to 20.41 US dollars per 100 km. The optimal sizes and costs for different price conditions are summarized in Table 7.

Figure 16 gives the comparative results of optimal costs under three price conditions. Based on the price trends in future, the optimal cost gradually decreases from 2020 to 2030. Compared to 2020, the decreases have reached 14.4%

and 18.59% in 2025 and 2030, respectively. The main reason is the sharp decrease of fuel cell and battery price. At the same time, benefiting from the decrease of fuel cell and battery price, the proportion of lifetime cost decreases rapidly. For example, based on the price performance in 2020, the proportion of lifetime cost is 25.56%, and it decreases to 18.26% and 14.35% in 2025 and 2030, respectively. Correspondingly, the hydrogen costs are 74.44%, 81.74%, and 85.65% for the price performance in 2020, 2025, and 2030, respectively. Along with the decreasing of hydrogen and components price, the hydrogen cost will be the more important performance index for vehicle cost in the near future.

7. Conclusion

This paper presents a sizing methodology for a heavy-duty fuel cell commercial vehicle on the basis of long-term stochastic drive cycle, parameters' scanning model, and DP-based EMS. The causes of hydrogen consumption under different sizes are studied in detail, and a discussion of the vehicle cost in the next decade is also given, on the basis of propulsion system and hydrogen price performance.

The main findings are as follows:

- (1) The hydrogen consumption will be minimized when the peak power of fuel cell stack reaches 120 kW, and the global propulsion efficiency will be improving with the battery cell number increase.
- (2) When the number of battery cells is small, the global propulsion efficiency is affected by the braking-recovery-energy and the fuel cell stack efficiency, while as the number of battery cells continues to increase, the main cause for that is the fuel cell stack efficiency. Limited by the maximum efficiency of fuel cell stack, the effect of increasing the battery cells number on the global propulsion efficiency will gradually decrease.
- (3) The size of $P_{\text{stack_max}} = 70$ kW and $n_{\text{bat}} = 696$ will be a better choice for the vehicle total cost, under the consideration of the cost performance in 2020. The predictive analysis shows that the larger fuel cell stack will be suitable for the FCHEV with the decrease of fuel cell and battery price.

Data Availability

The datasets and codes of this paper for the simulation are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Authors' Contributions

Weiguang Zheng carried out conceptualization, formal analysis, writing, review, and editing. Weiwei Xin was

responsible for software and wrote the original draft. Enyong Xu carried out conceptualization and data curation. Shuilong He carried out conceptualization and formal analysis. Jirong Qin carried out supervision and project administration. Heng Wang carried out methodology and validation. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This study was supported by Innovation-Driven Development Special Fund Project of Guangxi (Guike AA18242033), Liuzhou Science Research and Planning Development Project (nos. 2020GAAA0403 and 2019AD10203), Liudong Science and Technology Project (20200108), and Innovation Project of GUET Graduate Education (2019YCXS008).

References

- [1] L. Zhang, Y. Wang, and Z. Wang, "Robust lateral motion control for in-wheel-motor-drive electric vehicles with network induced delays," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10585–10593, 2019.
- [2] X. Ding, Z. Wang, and L. Zhang, "Hybrid control-based acceleration slip regulation for four-wheel-independent-actuated electric vehicles," *IEEE Transactions on Transportation Electrification*, vol. 7, pp. 1976–1989, 2020.
- [3] L. Guzzella and A. Sciarretta, *Vehicle Propulsion Systems: Introduction to Modeling and Optimization*, Springer, Heidelberg, Germany, 3 edition, 2013.
- [4] X. Lv, Y. Qu, Y. Wang, C. Qin, and G. Liu, "A comprehensive review on hybrid power system for PEMFC-HEV: issues and strategies," *Energy Conversion and Management*, vol. 171, pp. 1273–1291, 2018.
- [5] Y. Kim, M. Figueroa-Santos, N. Prakash, S. B. Stanley, J. Siegel, and M. Rizzo, "Co-optimization of speed trajectory and power management for a fuel-cell/battery electric vehicle," *Applied Energy*, vol. 260, Article ID 114254, 2020.
- [6] F. Millo, J. Zhao, L. Rolando, C. Cubito, and R. Fuso, "Optimizing the design of a plug-in hybrid electric vehicle from the early phase: an advanced sizing methodology," *Computer-Aided Design and Applications*, vol. 12, no. 1, pp. 22–32, 2015.
- [7] S. Onori, L. Serrao, and G. Rizzani, *Hybrid Electric Vehicles Energy Management Strategies*, Springer, New York, USA, 1st edition, 2016.
- [8] Y. Huang, H. Wang, A. Khajepour, H. He, and J. Ji, "Model predictive control power management strategies for HEVs: a review," *Journal of Power Sources*, vol. 341, pp. 91–106, 2017.
- [9] Y. Hames, K. Kaya, E. Baltacioglu, and A. Turksoy, "Analysis of the control strategies for fuel saving in the hydrogen fuel cell vehicles," *International Journal of Hydrogen Energy*, vol. 43, no. 23, pp. 10810–10821, 2018.
- [10] W. Zhou, L. Yang, Y. Cai, and T. Ying, "Dynamic programming for new energy vehicles based on their work modes Part II: fuel cell electric vehicles," *Journal of Power Sources*, vol. 407, pp. 92–104, 2018.
- [11] X. Li, Y. Wang, D. Yang, and Z. Chen, "Adaptive energy management strategy for fuel cell/battery hybrid vehicles using Pontryagin's Minimal Principle," *Journal of Power Sources*, vol. 440, Article ID 227105, 2019.
- [12] X. Lv, Y. Wu, J. Lian et al., "Energy management of hybrid electric vehicles: a review of energy optimization of fuel cell hybrid power system based on genetic algorithm," *Energy*

- Conversion and Management*, vol. 205, Article ID 112474, 2020.
- [13] G. Paganelli, S. Delpra, T. M. Guerra, J. Rimaux, and J. J. Santin, "Equivalent consumption minimization strategy for parallel hybrid Powertrains," in *Proceedings of the Vehicular Technology Conference*, pp. 2076–2081, Birmingham, AL, USA, May 2002.
- [14] H. Borhan, A. Vahidi, A. M. Phillips, M. L. Kuang, I. V. Kolmanovsky, and S. Di Cairano, "MPC-based energy management of a power-split hybrid electric vehicle," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 593–603, 2012.
- [15] Z. Chen, N. Guo, E. Xiao, J. Shen, and P. Dong, "A hierarchical energy management strategy for power-split plug-in hybrid electric vehicles considering velocity prediction," *IEEE Access*, vol. 6, Article ID 33261, 2017.
- [16] N. Guo, X. Zhang, Y. Zou, L. Guo, and G. Du, "Real-time predictive energy management of plug-in hybrid electric vehicles for coordination of fuel economy and battery degradation," *Energy*, vol. 214, Article ID 119070, 2021.
- [17] Z. Hu, J. Li, L. Song et al., "Multi-objective energy management optimization and parameter sizing for proton exchange membrane hybrid fuel cell vehicles," *Energy Conversion and Management*, vol. 129, pp. 108–121, 2016.
- [18] S. M. Lukic and A. Emadi, "Effects of drivetrain hybridization on fuel economy and dynamic performance of parallel hybrid electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 2, pp. 385–389, 2004.
- [19] M.-J. Kim and H. Peng, "Power management and design optimization of fuel cell/battery hybrid vehicles," *Journal of Power Sources*, vol. 165, no. 2, pp. 819–832, 2007.
- [20] D. Feroldi and M. Carignano, "Sizing for fuel cell/supercapacitor hybrid vehicles based on stochastic driving cycles," *Applied Energy*, vol. 183, pp. 645–658, 2016.
- [21] S. H. Karaki, R. Dinnawi, R. Jabr, R. Chedid, and F. Panik, "Fuel cell hybrid electric vehicle sizing using ordinal optimization," *SAE International Journal of Passenger Cars—Electronic and Electrical Systems*, vol. 8, no. 1, pp. 60–69, 2015.
- [22] X. Hu, N. Murgovski, L. Johannesson, and B. Egardt, "Energy efficiency analysis of a series plug-in hybrid electric bus with different energy management strategies and battery sizes," *Applied Energy*, vol. 111, pp. 1001–1009, 2013.
- [23] X. Hu, L. M. Johannesson, and B. Egardt, "Optimal dimensioning and power management of a fuel cell/battery hybrid bus via convex programming," *IEEE/ASME transactions on mechatronics*, vol. 20, pp. 457–468, 2014.
- [24] X. Hu, N. Murgovski, L. M. Johannesson, and B. Egardt, "Comparison of three electrochemical energy buffers applied to a hybrid bus powertrain with simultaneous optimal sizing and energy management," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 3, pp. 1193–1205, 2014.
- [25] X. Hu, S. J. Moura, N. Murgovski, B. Egardt, and D. Cao, "Integrated optimization of battery sizing, charging, and power management in plug-in hybrid electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 24, pp. 1036–1043, 2015.
- [26] M. H. Karaoglan, N. S. Kuralay, and Z. Colpan, "The effect of gear ratios on the exhaust emissions and fuel consumption of a parallel hybrid vehicle powertrain," *Journal of Cleaner Production*, vol. 210, pp. 1033–1041, 2018.
- [27] T. Markel, A. Brooker, T. Hendricks et al., "ADVISOR: a systems analysis tool for advanced vehicle modeling," *Journal of Power Sources*, vol. 110, no. 2, pp. 255–266, 2002.
- [28] A. Sciarretta and L. Guzzella, "Control of hybrid electric vehicle," *IEEE Control Systems Magazine*, vol. 27, pp. 60–70, 2007.
- [29] O. Sundstrom and L. Guzzella, "A generic dynamic programming Matlab function," in *Proceedings of the 18TH IEEE International Conference Conference on Control Applications*, Saint Petersburg UK Russia, July 2009.
- [30] O. Sundström, D. Ambühl, and L. Guzzella, "On implementation of dynamic programming for optimal control problems with final state constraints," *Oil & Gas Science and Technology - Revue de l'Institut Français du Pétrole*, vol. 65, no. 1, pp. 91–102, 2010.
- [31] E. T. Yidiz, Q. Farooqi, S. Anwar, Y. Chen, and A. Izadian, "Nonlinear constrained component optimization of a plug-in hybrid electric vehicle powertrain," in *Proceedings of the 25th World Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exhibition*, Shenzhen, China, September 2010.
- [32] H. Li, Y. Zhou, H. Gualous, H. Chaoui, and L. Boulon, "Optimal cost minimization strategy for fuel cell hybrid electric vehicles based on decision-making framework," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2388–2399, 2021.

Research Article

Solving No-Wait Flow Shop Scheduling Problem Based on Discrete Wolf Pack Algorithm

Rongshen Lai ^{1,2}, Bo Gao,¹ and Wenguang Lin ^{1,2}

¹School of Mechanical and Automotive Engineering, Xiamen University of Technology, Xiamen 361024, China

²Xiamen Key Laboratory of Intelligent Manufacturing Equipment, Xiamen 361024, China

Correspondence should be addressed to Rongshen Lai; 122774584@qq.com

Received 5 November 2021; Accepted 6 December 2021; Published 21 December 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Rongshen Lai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the no-wait flow shop scheduling problem with the goal of minimizing the maximum makespan, a discrete wolf pack algorithm has been proposed. First, the methods for solving the no-wait flow shop scheduling problem and the application research of the wolf pack algorithm were summarized, and it was pointed out that there was lack of research on the application of the wolf pack algorithm to solve the no-wait flow shop scheduling problem. According to the analysis of characteristics of the no-wait flow shop scheduling problem, the individual wolf was coded by a decimal integer; wolf searching behavior was realized through the exchange of different code bits in the individual wolf, and the continuous code segment of the head wolf was randomly selected to replace the corresponding code of the fierce wolf, by which the behaviors of wolves raiding and sieging were realized, and the population was updated according to the rule of “survival of the strong.” In particular, to fully explore the potential optimal solution in the solution space, loop operations were added to the wandering, summoning, and siege processes. Finally, based on a comparison with the leapfrog algorithm and the genetic algorithm, the effectiveness of the algorithm was verified.

1. Introduction

Production scheduling is a key link to ensure the efficient and orderly development of the manufacturing process. It is an important way to quickly respond to customer needs, improve corporate economic efficiency, and maintain market competitiveness. Research on scheduling issues has important theoretical and practical significance in the current intelligent manufacturing context. The no-wait flow shop scheduling problem (NWFSP) is a very important type of scheduling problem, which widely exists in food processing, chemical, metallurgy, and pharmaceutical industries, and is also a typical NP-hard problem [1]. Based on the inspiration of the marvelous group phenomenon in nature, researchers have now proposed many effective swarm intelligence optimization algorithms to solve this problem, such as genetic algorithm [2], particle swarm algorithm [3], ant colony algorithm [4], etc. The development of swarm intelligence optimization algorithms is in the ascendant, providing many options for solving complex optimization

problems. The wolf pack algorithm (WPA) is a group intelligence optimization algorithm obtained by simulating the hunting activities of the wolf pack. It has the advantages of strong global search ability, good generalization ability, and easy operation. It has significant effect on processing multi-peak and high-dimensional complex functions, especially, WPA is suitable for solving various complex combinatorial optimization problems, such as TSP and knapsack problem. [5]. However, currently, there are few research reports on the application of the wolf pack algorithm to the NWFSP. To this end, this paper combines the implementation process of the wolf pack algorithm and the feature analysis of NWFSP and proposes an improved discrete wolf pack algorithm and proves its effectiveness through a practical example and comparison with the leapfrog algorithm (LFA) and the genetic algorithm (GA).

The remainder of this paper is organized as follows. In Section 2, a state-of-the-art about method for solving the no-wait flow shop scheduling problem and wolf pack algorithm is provided. In Section 3, the mathematical model of the no-

wait flow shop scheduling problem is described. In Section 4, the proposed improved discrete wolf pack algorithm and the corresponding algorithm process are illustrated in detail. In Section 5, a case study and comparison with particle swarm algorithm and genetic algorithm are carried out to verify the effectiveness of the proposed algorithm. Finally, the conclusion and future research directions are pointed out in Section 6.

2. Research Status

Focusing on the research of discrete wolf pack algorithm for solving no-wait flow shop scheduling problem, this section focuses on the literature review from two aspects: the method for solving the no-wait flow shop scheduling problem and the application of wolf pack algorithm.

2.1. Research on Methods for Solving No-Wait Flow Shop Scheduling Problem. The no-wait flow shop scheduling model is mainly aimed at the production processes that cannot be interrupted, such as steel rolling, food production, and so on. This model has a wide range of applications and is difficult to solve. It has attracted the attention and in-depth research of many experts and scholars, such as Song et al. [6] proposed a neighborhood iterative search algorithm for NWFSP, which reduced the time complexity of the solution process and enhanced the ability to find the global optimal solution. Zhang and Yu [7] aimed at the NWFSP with makespan minimization and proposed a discrete fruit fly optimization algorithm based on the dominant population. Orhan and Abdullah [8] aimed at the non-wait flow shop scheduling problem with makespan minimization as the criterion, and proposed a new hybrid ant colony algorithm based on crossover and mutation mechanism, and the performance of the algorithm was compared with adaptive learning methods and genetic heuristic algorithms; Zhao et al. [9] used hybrid biogeographic optimization algorithm and variable neighborhood search algorithm comprehensively to solve NWFSP. Allahverdi [10] carried out a systematic review of no-wait flow scheduling problems.

2.2. Research on Wolf Pack Algorithm. Wolves are a highly social species with a strict hierarchy and strong domain awareness. Wolves are usually led by the head wolves with absolute superiority. They kill the prey through a clear division of labor and cooperation among members and distribute food according to the “survival of the strong” rule. Liu et al. [11] simulated the intelligent hunting behavior of wolves, abstracted the three behaviors of searching for prey, besieging prey, and updating wolves, and proposed the wolf colony algorithm (WCA) in 2011 to solve the optimization problem. The main processes included assigning artificial wolves from wolves to search prey within the range of prey activities. Once the prey was found, other artificial wolves will be notified of the position of the prey by howling, and other artificial wolves will approach the prey to encircle. The WCA mainly included five steps: initialization, selection of wolves to detect prey, treating the optimal position of some

artificial wolf as the position of the prey, updating the wolf pack according to the “survival of the strong” rule, and judging whether the termination condition was met.

Based on the analysis of the characteristics of wolf pack cooperative hunting and prey distribution, Wu et al. [5] abstracted 3 kinds of artificial wolves (head wolf, searching wolf, fierce wolf), 3 kinds of intelligent behaviors (wandering, summoning, siege), and 2 kinds of intelligent rules (wolf generation rule “winner is king” and wolf pack update mechanism “survival of the strong”), and proposed the wolf pack algorithm (WPA) with different optimization strategy compared with WCA in 2013. The convergence of the algorithm was proved based on Markov chain theory, and the comparison with other algorithms verified that the algorithm had better global convergence and computational robustness. Based on WPA, Hui et al. [12] proposed an improved wolf pack algorithm in 2017 by the introduction of the concept of siege radius and optimization of step length and design the position update formula of fierce wolves. Based on the research of grey wolf hunting behavior [13], Mirjalili et al. [14] proposed a new meta-inspiration algorithm—grey wolf algorithm (GWA) in 2014, which simulated the leadership hierarchy and the hunting mechanism of grey wolf groups in nature, abstracted 4 grey wolf levels and 3 main steps of hunting, namely searching, encircling, and attacking prey, and finally compared with particle swarm optimization and some meta-heuristic algorithms such as gravity search, differential evolution, evolution planning and evolution strategy to verify the effectiveness of the algorithm.

The wolf pack algorithm has good performance in global search and local development capabilities. Since its proposal, it has continuously attracted the attention of scholars and has been quickly applied to engineering practice. Yi et al. [15] proposed a hierarchic wolf pack algorithm to solve the problem of optimal placement of sensors; Wu et al. used the wolf pack algorithm to solve the binary knapsack problem [16], the traveling salesman problem [17], and unconstrained global optimization problem [18]. Fang and Tang [19] used an improved wolf pack algorithm to solve the three-dimensional routing optimization problem for AVE/RS composite operation. Liu et al. [20] used the wolf pack algorithm to plan the UAV track with a known starting point and endpoint. Wang and Jiao [21] proposed an improved wolf pack algorithm to solve the optimal scheduling problem of hydropower stations and reservoirs. Xie and Zhang [22] proposed a discrete wolf pack algorithm according to the characteristics of the permutation flow shop scheduling problem.

It was found from the research review that compared with some other algorithms such as particle swarm algorithm, dynamic programming algorithm, et al., the wolf pack algorithm had shown stronger optimization ability and faster convergence speed in the process of solving a combinatorial optimization problem. However, aiming at the typical combinatorial optimization problem, the no-wait flow shop scheduling problem, the application research of the wolf pack algorithm was relatively lacking. To this end, an improved discrete wolf pack algorithm was proposed in this paper to solve the no-wait flow shop scheduling problem.

3. Mathematical Model of NWFSP

The no-wait flow shop scheduling problem can be described as follows [3].

Given: (1) m machines and n workpieces. (2) The processing sequence of the workpieces on the machines is the same. (3) The processing time of each workpiece on each machine. (4) All workpieces can be processed at zero time.

Constraints: (1) A workpiece can only be processed on one machine at a certain time. (2) A machine can only process one workpiece at a certain time. (3) The transportation time of the workpiece and the start-up time of the machine is included in processing time. (4) All processes of the same workpiece must be processed continuously, that is, once each workpiece starts to be processed, each process must be performed continuously, and there is no waiting time between two adjacent processes.

Goal: To determine a scheduling plan that minimizes the maximum makespan.

Based on the literature [23], assuming that the processing time of the workpiece i on the machine k is $p_{i,k}$, according to the continuous production requirement, the difference between the start time of two adjacent workpieces $i-1$ and i (start processing time interval) $d_{i-1,i}$ shall meet the following requirement, as shown in(1):

$$d_{i-1,i} = \max \left\{ \max_{2 \leq k \leq m} \left\{ \sum_{q=1}^k p_{i-1,q} - \sum_{q=1}^{k-1} p_{i,q} \right\}, p_{i-1,1} \right\}. \quad (1)$$

The maximum makespan is calculated as:

$$T_{\max} = \sum_{j=2}^n d_{j-1,j} + \sum_{k=1}^m p_{n,k}. \quad (2)$$

According to the analysis, the calculation complexity of the maximum makespan is $O(mn^2)$.

4. Improved Discrete Wolf Pack Algorithm

The wolf pack algorithm realized the whole process simulation of the searching of prey and environmental information detection by individual wolves, the sharing and interaction of information between artificial wolves, and the whole process of artificial wolf capturing prey based on individual behavior decisions of their own responsibilities. Wolf pack algorithm [5] consists of three intelligent behaviors of wandering, summoning, and besieging, and the “winner is king” rule of wolf competition, and the “survival of the strong” wolf pack update mechanism. Based on the discrete wolf pack algorithm for solving the TSP problem proposed by Wu et al. [17], in view of the characteristics of NWFSP, this section introduces the population initialization, intelligent behaviors, and rule description of the improved discrete wolf pack algorithm in detail.

4.1. Coding Rules and Population Initialization. NWFSP is a typical discrete combinatorial optimization problem. According to the characteristics of the problem, the decimal encoding method is adopted, that is, each workpiece is represented by a decimal integer, and all the workpieces are processed on the machine according to the predetermined process. The processing sequence of all the workpieces constitutes a decimal sequence $X_i = (x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{in})$, which is used to represent the position of the i^{th} artificial wolf in the wolf pack algorithm, where n represents the total number of workpieces to be processed, and x_{ij} represents the coding number of the j^{th} processed workpiece in the sequence X_i . Taking the scheduling problem of 7 workpieces (coded from 1 to 7) as an example, the decimal sequence $X_i = \{4, 6, 3, 2, 5, 1, 7\}$ indicates that the processing order of the 7 workpieces is workpiece 4 \rightarrow workpiece 6 \rightarrow workpiece 3 \rightarrow workpiece 2 \rightarrow workpiece 5 \rightarrow workpiece 1 \rightarrow workpiece 7.

Determining the population size to be N , and using the random generation method to get the initial population $X = \{X_1, X_2, \dots, X_i, \dots, X_N\}$, where $1 \leq i \leq N$.

4.2. Intelligent Behavior and Rules. Based on the analysis of the characteristics of wolf pack cooperative hunting activities, the wolf pack algorithm abstracts the 3 intelligent behaviors of wolf pack wandering, summoning, and besieging, combined with the solution goal of NWFSP (that is, minimizing the maximum makespan), and elaborates the corresponding behavior rules in detail.

4.2.1. Selection of Head Wolf. The artificial wolf with the optimal objective function value (that is, the shortest makespan) in the initial population is selected as the head wolf; in the iteration process, after each iteration, the objective function value of the current head wolf is compared with the objective function values of other artificial wolves, If there is an artificial wolf whose objective function value is better than that of the head wolf, that artificial wolf is used to replace the head wolf; if the optimal objective function value corresponds to multiple artificial wolves, one of them is randomly selected as the head wolf.

4.2.2. Wandering Behavior. All artificial wolves in the wolf pack except the head wolf are regarded as detecting wolves to search for prey in the solution space. Assume that Y_i^0 and Y_{lead} respectively represent the prey odour concentration perceived by the wolf i and the head wolf in the initial population. The maximum number of wanderings are set as K . During the first wandering process, let the wolf i take one step forward in h directions respectively (the step length at this time is called the wandering step length $step_a$), and record the position in the p^{th} direction and the perceived prey odour concentration Y_i^p ($1 \leq p \leq h$), then return to the original position; if $Y_i^p \geq Y_{\text{lead}}$, the detecting wolf i will replace the head wolf; if $Y_i^0 < Y_i^p < Y_{\text{lead}}$, use the coordinates of one step forward in the p^{th} direction to replace the coordinates of the detecting wolf i before wandering;

if $Y_i^p < Y_i^0 < Y_{lead}$, the coordinates of detecting wolf i remain unchanged. Specifically, the detecting wolf i takes one step forward in the p -th direction ($p = 1, 2, \dots, h$), that is, randomly interchange two selected workpieces x_{ij} and x_{ik} in the code $X_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}$ of the detecting wolf i . Or randomly select a workpiece from the code $X_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}$ of the detecting wolf i , and then insert the workpiece after other workpieces from left to right. Repeat the above-mentioned wandering behavior until the maximum number of wanderings is reached, and then proceed to the summoning phase of the head wolf.

4.2.3. Summoning Behavior. All artificial wolves in the pack except the head wolf are regarded as fierce wolves. When the head wolf howls to summon the fierce wolves, the fierce wolves quickly rush towards the current position of the head wolf with a large step length. Referring to related literature [11], the summoning behavior is designed as follows.

Randomly select a piece of continuous code $x_{is}, x_{i(s+1)}, \dots, x_{ie}$ with the starting point x_{is} and the ending point x_{ie} in the head wolf, respectively, where the number of code digits is the raiding step length $step_b$, replace a segment of continuous code $x_{is}, x_{i(s+1)}, \dots, x_{ie}$ in the corresponding position of the fierce wolf i . This operation reflects the leadership of the head wolf (the optimal individual) to the wolf pack. If the elements of these two continuous codes are the same, the codes in other positions in the fierce wolf i will not be changed, which reflects the differentiation of the wolf pack individual. The element in $x_{is}, x_{i(s+1)}, \dots, x_{ie}$ that does not belong to $x_{is}, x_{i(s+1)}, \dots, x_{ie}$ are randomly placed in the rest of the fierce wolf i . The specific operation is as follows: Suppose the code of head wolf is $X_{lead} = \{5, 3, 4, 2, 8, 1, 7, 6\}$, and the code of the fierce wolf i is $X_i = \{4, 6, 3, 2, 8, 5, 7, 1\}$, randomly select a piece of continuous code from the head wolf as $\{4, 2, 8, 1\}$, and the corresponding code in the fierce wolf i is $\{3, 2, 8, 5\}$; use the code segment $\{4, 2, 8, 1\}$ to replace the code segment $\{3, 2, 8, 5\}$, then the code $\{4, 6, 4, 2, 8, 1, 7, 1\}$ is obtained. In this code, the elements 4 and 1 appear repeatedly, and the elements 3 and 5 are missing. The elements 6 and 7 in the fierce wolf i does not belong to $\{4, 2, 8, 1\}$, so keep the position unchanged; randomly arrange 3 and 5 and replace the repeated elements 4 and 1, and finally get the updated code $\{3, 6, 4, 2, 8, 1, 7, 5\}$ or $\{5, 6, 4, 2, 8, 1, 7, 3\}$ of the fierce wolf i .

After updating the code of the fierce wolf i , recalculate its fitness value. If $Y_i > Y_{lead}$, the fierce wolf i will replace the head wolf. After reaching the set number of raids, the process enters the siege behavior.

4.2.4. Siege Behavior. After the raid process, the distance between the wolves and the prey is relatively close, and the head wolf will command the fierce wolves to besiege the prey. The design of the siege behavior is similar to the summoning behavior. In order to ensure that the wolves perform a fine search near the prey, the siege step length $step_c$ should not be greater than the raiding step length $step_b$ at this time, and the siege behavior ends when the set number of siege is reached.

4.2.5. Population Update. In order to prevent the population from entering the local optimum, the population is updated after each iteration. The specific operation is as follows: arrange the population from large to small according to the fitness value, remove the last R artificial wolves in the population, and then randomly generate R artificial wolves to join the population for the next iteration.

4.3. Algorithm Flow. The specific process of the improved discrete wolf pack algorithm is shown in Figure 1 and described in detail as follows.

Step 1: Parameter initialization. Set the population size of the wolf pack N , the maximum number of wandering times T_{max} , the maximum number of iterations of the algorithm k_{max} , the range of the search direction h , and the number of updates of the wolf pack R .

Step 2: Initialize the spatial position of the wolf pack, calculate its objective function value, and select the artificial wolf with the optimal fitness value as the head wolf X_{lead} .

Step 3: Wolf pack detecting. If the function value of the detecting wolf is greater than the head wolf, it will replace the head wolf and initiate a summoning behavior. Otherwise, the detecting wolves will continue to wander until the maximum number of wanderings is reached, the head wolf will summon the other wolves.

Step 4: Wolf pack raiding. The fierce wolves rush towards the prey. If the prey odour concentration of some fierce wolf is greater than that of the head wolf, this fierce wolf will replace the head wolf, and the fierce wolves will rush to the range close to the prey, and proceed to the next step.

Step 5: Wolf pack siege. If the prey odour concentration of some fierce wolf is greater than that of the head wolf, this fierce wolf will replace the head wolf.

Step 6: Population renewal. Update the position of the head wolf according to the "winner is king" rule, update the wolf pack according to the "survival of the strong" mechanism, and then enter the next iteration process.

Step 7: Determine whether the accuracy requirement is met or the maximum number of iterations is reached. If the algorithm termination condition is met, the optimal solution is output, that is, the position of the head wolf or the position of the prey; otherwise, go to Step 3 to continue.

5. Performance Verification

In order to verify the feasibility of the improved discrete wolf pack algorithm (IDWPA) designed in this paper in solving the NWFSP problem, this paper uses 5 sets of calculation examples to compare the algorithm with LFA and GA [24]. The termination condition of the algorithm is reaching the number of iterations. The initialization parameters of IDWPA are set as follows: the population size of the wolf

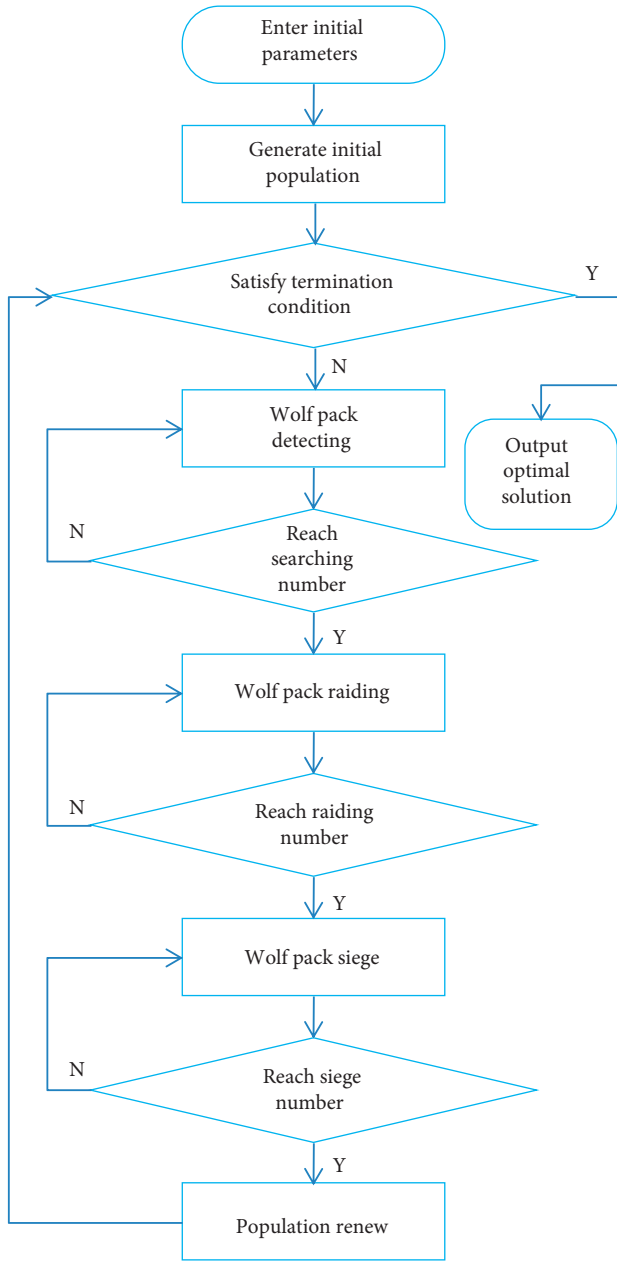


FIGURE 1: Flowchart of the improved discrete wolf pack algorithm.

pack $N=20$, the maximum number of wandering times $T_{max}=10$, the maximum number of iterations of the algorithm $k_{max}=100$, the range of the search direction $h=5$, and the number of updates of the wolf pack $R=4$.

Table 1 shows the optimal solution and the average solution with three different algorithms. Compared with LFA and GA, IDWPA designed in this paper reduces the makespan of the NWFSP problem and improves the average equipment utilization rate.

It can be seen from the population evolution iterative process that the improved discrete wolf pack algorithm is easier to jump out of the local optimum and has a faster convergence rate when solving the NWFSP problem. Furthermore, the scheduling result is more accurate, and the utilization rate of the machine is improved to a certain

TABLE 1: Scheduling solutions of three algorithms.

Case	Optimal solution			Average solution		
	IDWPA	LFA	GA	IDWPA	LFA	GA
10×10	1249	1249	1249	1262	1266	1285
15×15	1708	1719	1756	1742	1743	1808
20×10	1756	1756	1766	1776	1782	1877
20×20	1220	1217	1228	1287	1291	1298
50×10	3545	3509	3555	3580	3584	3654

extent, so the production efficiency of the enterprise can be improved to a certain extent.

6. Conclusions

Aiming at the no-wait flow shop scheduling problem, an improved discrete wolf pack algorithm was proposed. First of all, this paper summarized the method of solving the problem of no-waiting flow shop scheduling and the research on the application of the wolf pack algorithm and points out the research gap of using the wolf pack algorithm to solve NWFSP. Then, the decimal integer coding method was adopted; the wandering behavior of wolf detecting was realized through the exchange of single code bits, the raiding and siege behaviors were realized by replacing the continuous code segment. The population was updated according to the rule of “survival of the strong,” and the summoning and siege links were added with cyclic operations. Finally, a comparison with the leapfrog algorithm and genetic algorithm was performed to verify the effectiveness of the proposed algorithm.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

All the authors do not have any possible conflicts of interest.

Acknowledgments

This research was supported by the Xiamen Key Laboratory of Intelligent Manufacturing Equipment, Fujian Province Education and Research Project for Young and Middle-Aged Teacher (Grant nos. JAT200467 and JAT200472), Fujian Social Science Fund (FJ2021B128), the Fujian Supporting Project of STS Plan of Chinese Academy of Sciences (2021T3069), and the National Innovation Method Fund of China (No. 2019IM010300).

References

- [1] N. G. Hall and C. Sriskandarajah, “A survey of machine scheduling problems with blocking and no-wait in process,” *Operations Research*, vol. 44, no. 3, pp. 510–525, 1996.
- [2] X. Zhu, X. Li, and Q. Wang, “Based on the total idle time increment no-waiting flow scheduling hybrid genetic

- algorithm,” *Computer Research and Development*, vol. 48, no. 3, pp. 455–463, 2011.
- [3] Q. Pan, W. Wang, and J. Zhu, “Modified discrete particle swarm optimization algorithm for no-wait flow shop problem,” *Computer Integrated Manufacturing Systems*, vol. 13, no. 6, pp. 1127–1130+1136, 2007.
 - [4] Q. Pan, B. Zhao, Y. Qu, and Y. Bi, “Ant-colony heuristic algorithm for no-wait flow shop problem with makespan criterion,” *Computer Integrated Manufacturing Systems*, vol. 13, no. 9, pp. 1801–1804+1815, 2007.
 - [5] H. Wu, F. Zhang, and L. Wu, “New swarm intelligence algorithm-wolf pack algorithm,” *Systems Engineering and Electronics*, vol. 35, no. 11, pp. 2430–2438, 2013.
 - [6] C. Song, X. Liu, and W. Wang, “An iterative neighbourhood search algorithm for large scale no-wait flow-shop,” *Control and Decision*, vol. 26, no. 4, pp. 535–539, 2011.
 - [7] Q. Zhang and Z. Yu, “Discrete fruit fly optimization algorithm based on dominant population for solving no-wait flow shop scheduling problem,” *Computer Integrated Manufacturing Systems*, vol. 23, no. 3, pp. 609–615, 2017.
 - [8] E. Orhan and G. Abdullah, “A new hybrid ant colony optimization algorithm for solving the no-wait flow shop scheduling problems,” *Applied Soft Computing*, vol. 72, pp. 166–176, 2018.
 - [9] F. Zhao, S. Qin, Y. Zhang, W. Ma, C. Zhang, and H. Song, “A hybrid biogeography-based optimization with variable neighborhood search mechanism for no-wait flow shop scheduling problem,” *Expert Systems with Applications*, vol. 126, no. JUL, pp. 321–339, 2019.
 - [10] A. Allahverdi, “A survey of scheduling problems with no-wait in process,” *European Journal of Operational Research*, vol. 255, no. 3, pp. 665–686, 2016.
 - [11] C. Liu, X. Yan, and H. Wu, “The wolf colony algorithm and its application,” *Chinese Journal of Electronics*, vol. 20, no. 2, pp. 212–216, 2011.
 - [12] X. Hui, Q. Guo, P. Wu et al., “An improved wolf pack algorithm,” *Control and Decision*, vol. 32, no. 7, pp. 1163–1172, 2017.
 - [13] C. Muro, R. Escobedo, L. Spector, and R. P. Coppinger, “Wolf-pack (*Canis lupus*) hunting strategies emerge from simple rules in computational simulations,” *Behavioural Processes*, vol. 88, no. 3, pp. 192–197, 2011.
 - [14] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
 - [15] T. Yi, C. Wang, and H. Li, “Hierarchic wolf algorithm for optimal triaxial sensor placement,” *Journal of Building Structures*, vol. 35, no. 4, pp. 223–229, 2014.
 - [16] H. Wu, F. Zhang, R. Zhan, and S. Wang, “A binary wolf pack algorithm for solving 0-1 knapsack problem,” *Systems Engineering and Electronics*, vol. 36, no. 8, pp. 1660–1667, 2014.
 - [17] H. Wu, F. Zhang, H. Li et al., “Discrete wolf pack algorithm for traveling salesman problem,” *Control and Decision*, vol. 30, no. 10, pp. 1861–1867, 2015.
 - [18] H.-S. Wu and F.-M. Zhang, “Wolf pack algorithm for unconstrained global optimization,” *Mathematical Problems in Engineering*, vol. 2014, no. 1, pp. 1–17, Article ID 46508, 2014.
 - [19] Y. Fang and M. Tang, “Three-dimensional routing optimization for AVS/RS’s composite operation,” *Computer Integrated Manufacturing Systems*, vol. 21, no. 3, pp. 702–708, 2015.
 - [20] Y. Liu, W. Li, H. Wu et al., “Track planning for unmanned aerial vehicles based on wolf pack algorithm,” *Journal of System Simulation*, vol. 27, no. 8, pp. 1838–1843, 2015.
 - [21] J. Wang and Y. Jiao, “Improvement of wolf pack algorithm and its application to optimal operation of reservoirs,” *Engineering Journal of Wuhan University*, vol. 50, no. 2, pp. 161–167+173, 2017.
 - [22] R. Xie and H. Zhang, “Discrete wolf pack algorithm for permutation flow shop scheduling problem,” *Control Engineering China*, vol. 27, no. 2, pp. 288–296, 2020.
 - [23] J. Grabowski and J. Pempera, “Some local search algorithms for no-wait flow-shop problem with makespan criterion,” *Computers & Operations Research*, vol. 32, no. 8, pp. 2197–2212, 2005.
 - [24] Y. Wu, Y. Wang, H. Zhang et al., “Improved frog leaping algorithm for solving no-waiting flow shop scheduling problem,” *Modular Machine Tool & Automatic Manufacturing Technique*, vol. 7, pp. 81–84, 2020.

Research Article

Investor Sentiment Combined with Multisource Information to Predict Stock Prices: An Analysis of China's A-Share Market

Xin Huang ¹ and Huilin Song ²

¹*School of Software, Jiangxi Normal University, Nanchang 330031, China*

²*School of International Economics and Trade, Jiangxi University of Finance and Economics, Nanchang 330013, China*

Correspondence should be addressed to Huilin Song; songhuilin@jxufe.edu.cn

Received 13 August 2021; Revised 20 October 2021; Accepted 22 November 2021; Published 7 December 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Xin Huang and Huilin Song. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Investor sentiment has been widely used in the research of the stock market, and how to accurately measure investor sentiment is still being explored. With the rise of social media, investor sentiment is no longer only influenced by macroeconomic data and news media, but also guided by We-Media and fragmented information. We take the data of China A-shares from January 2020 to December 2020 as the research object and propose a stock price prediction method that combines investor sentiment with multisource information. Firstly, the sentiment of macroeconomic data, brokerage research reports, news, and We-Media is calculated, respectively, and then the investor sentiment vector combining multisource information is obtained by the multilayer perceptron. Finally, the LSTM model is used to represent the stock time series characteristics. The results show that (1) the proposed algorithm is superior to the benchmark algorithm in terms of accuracy and F1-score, (2) investor sentiment vector can effectively measure the investment sentiment of stocks, and (3) compared with vector concatenation, multilayer perceptron can better represent investor sentiment.

1. Introduction

Behavioral finance, which is derived from finance, psychology, communication, and behavioral science, believes that the stock price is not only determined by the intrinsic value of an enterprise but is largely influenced by the psychology and behavior of investors [1]. The idea in behavioral finance is that investors in markets are not completely rational people. In the process of investment decision-making, investors often cannot make a correct and reasonable judgment due to factors such as emotional preference and cognitive bias. In other words, investor sentiment reflects investor behavior and affects the final investment decision to some extent. Researchers try to explain market behavior from the perspective of investors. To verify the effectiveness of investor sentiment, Akerlof and Shiller [2] found a close relationship between investor sentiment and stock price by studying the volatility of investor sentiment and stock price. You and Wu [3] used the “spiral of silence” theory in media

effect research of communication to study the impact of sentiment index on stock asset pricing from the perspective of media. Sentiment, as a factor affecting investors' psychological activities and then their behavior, has gradually become an important research issue in the task of stock price prediction.

Investor sentiment plays an important role in stock price forecasting. Song et al. [4] proposed a method for predicting stock excess returns that integrates research reports and investor sentiment, which can be verified in the Chinese A-share market to effectively improve the accuracy of the forecast. Li et al. [5] also conducted a similar study, and the difference is that the research object is the Hong Kong stock market. Polk and Sapienza [6] showed in their research that investor sentiment is similar to mispricing behavior in the stock market. Other views believe that investor sentiment is formed by the wrong estimation of asset value, which to a certain extent indicates the speculative propensity of investors [7]. Although the definition of investor sentiment

has not yet reached a unified concept, it can be seen from different definitions that investor sentiment is an expectation of future stock returns, and due to investors' irrational behavior and reasons that are not completely based on fundamental analysis, investors will have certain deviations in their expectations [8].

In the current study, measures of investor sentiment can be divided into three categories. The first is the direct measurement method, which uses indicators obtained from market surveys to directly replace investor sentiment. The second is the indirect measurement method, which uses single economic variables and combination variables as proxy variables to measure investor sentiment. The third type uses the machine learning method to extract online text information in social media and further construct an investor sentiment index to measure investor sentiment. The information explosion and fragmented nature of the age of big data make it inadequate to use any of these measures alone to measure investor sentiment. In our opinion, the measurement of investor sentiment should take into account four factors simultaneously: macroeconomic conditions, brokerage research reports, news, and We-Media information. Based on this, we put forward a kind of multisource information fusion method to predict the price of the shares of investor's emotion; first, the sentiment of macroeconomic data, securities research reports, news, and the media is calculated, fusion of multisource information is obtained by concatenation operation ISV (Investor Sentiment Vector, ISV), and finally LSTM model is used to represent the stock time series characteristics. The contribution of this paper is as follows:

- (1) An investor sentiment measurement method integrating multisource information is proposed
- (2) The positive role of investor sentiment in the stock prediction task is verified
- (3) A stock price prediction framework is proposed based on deep learning

The rest of this paper is organized as follows. Section 2 reviews investor sentiment measurement and its relationship with stock prices. Section 3 introduces our proposed method. Section 4 presents the experimental and details. Section 5 presents the experimental results and discussion. Section 6 gives our conclusions and directions for future work.

2. Related Works

With the continuous development of the Internet, the emergence of social media provides a new platform for users to search for information, express their feelings, and exchange opinions. Using social media indices as a proxy for investor sentiment has also become a convenient way to capture investor sentiment in the market. According to the Google search index, Da et al. [9] constructed the investor sentiment index through Google search keywords and found that the index could predict the short-term return and volatility of stocks. Meng et al. [10] use the Baidu search index to measure investor sentiment and find that investor

sentiment has a linkage mechanism with the stock market. Although quantitative indicators are feasible in reflecting investors' attention to the stock market, they are difficult to measure more in-depth investor sentiment information [11].

With the rise of big data, text mining, machine learning, and sentiment analysis technologies, researchers can more quickly and accurately extract valuable information from texts for the construction of investor sentiment [12]. Oliveira et al.'s [13] research shows that investor sentiment extracted from social media platforms has a certain impact on stock prices, and social media also provides a large number of data sources for the construction of investor sentiment. Bollen and Mao [14] analyzed and compared the predictive ability of traditional investor sentiment metrics and social media and found that sentiment indicators extracted from social media have a better predictive effect. Sentiment indicators obtained from text analysis of social media content have been widely used in stock market prediction, but there is no consistency in research conclusions [15]. Ma and Zhang [16] believed that the inconsistent conclusions were caused by the difference in sample data selection and the accuracy of investor sentiment measurement. At present, the research is no longer limited to judging whether investor sentiment can predict the stock market. How to extract valuable information from a large amount of data and apply it to the construction of investor index has become the focus of the research.

Pröllochs et al. [17] analyzed the information in financial news media and found that the sentiment of negative sentences in financial news is correlated with stock prices. In terms of information usefulness, Sprenger et al. [18] point out that many professional and amateur investors and analysts use Twitter to post news comments and opinions, usually more frequently than professional news media. In terms of the speed of information transmission, Sul et al. [19] believe that investors' emotions transmitted through social media are more likely to affect stock prices quickly, while investors' emotions that spread more slowly take longer to affect stock prices and are more likely to predict prices in the next few days. In addition to Twitter, StockTwits [20] and Yahoo Finance [21, 22] are also used to mine investor sentiment.

3. Method

We propose a stock price prediction method, including an investor sentiment module and a stock prediction module. The investor sentiment module separately calculates the four dimensions of macroeconomic status, broker report sentiment, news sentiment, and self-media sentiment through different methods and then obtains the ISV by MLP (multilayer perceptron) [23]. The stock prediction module consists of an LSTM [24], where the first input of LSTM is investor sentiment, and the subsequent input is the stock price. The method flow is shown in Figure 1.

3.1. Investor Sentiment Vector

3.1.1. Macroeconomic Status. MS (macroeconomic status) includes market status and economic status. For the market status, we select five indicators to measure the transaction

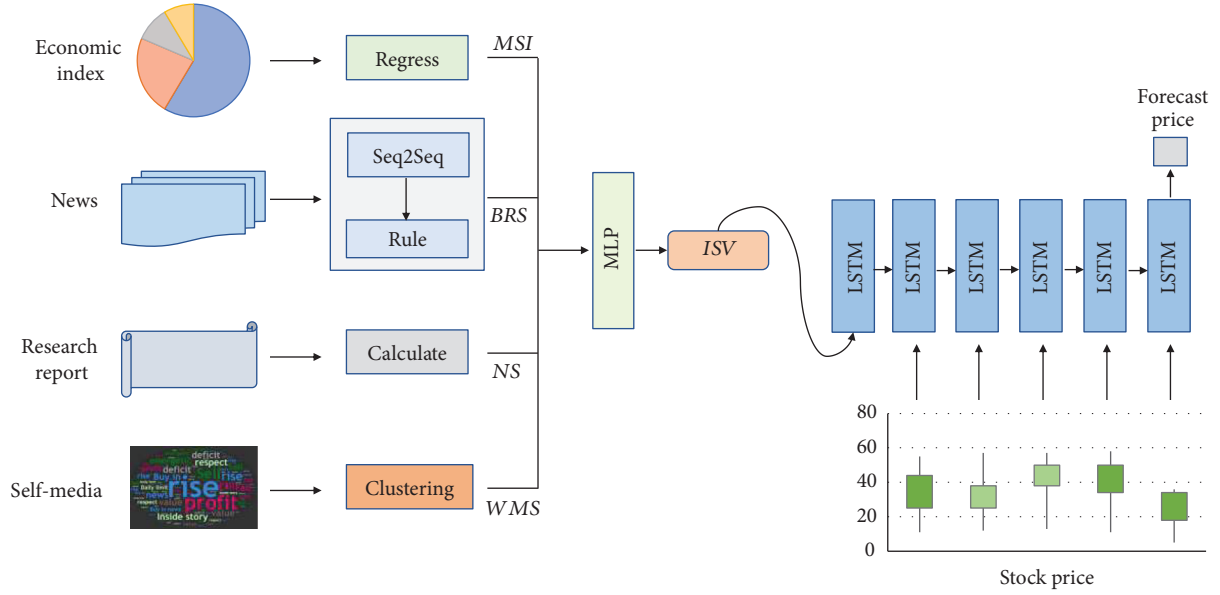


FIGURE 1: A stock price forecasting framework incorporating investor sentiment based on multiple-source information.

volume (VOLUME), the number of new investor accounts (NEWIN), the consumer confidence index (CCI), the closed-end fund discount rate (FUND), and the market turnover rate (HS_TVR). The economic status is measured by four indicators: the resident consumption index (CPI), the amount of new credit (IC), the economic growth rate (GDP), and the money supply (M2). The macroeconomic status measurement can be divided into two steps.

$$\text{Sentiment}^1 = a * \text{FUND} + b * \text{NEWIN} + c * \text{VOLUME} + d * \text{HS}_{\text{TVR}} + e * \text{CCI}. \quad (1)$$

The second step is to control the influence of economic state indicators and perform regression analysis on Sentiment^1_m , as shown in formula (2). The residual value ε_m is the measurement index of macrosentiment (MS).

$$\text{Sentiment}^1_m = \alpha_0 + \alpha_1 \text{CPI} + \alpha_2 \text{IC} + \alpha_3 \text{GDP} + \alpha_4 \text{M}_2 + \varepsilon_m. \quad (2)$$

Among them, m is the month, α_0 is a constant, and $\alpha_1 \sim \alpha_4$ are the regression coefficients to be estimated.

3.1.2. Brokerage Report Sentiment. Referring to the method proposed in the literature [4], we first split the brokerage research report into attention $\text{RRA}_{k,t}$ and rating sentiment $\text{RRRS}_{k,t}$ and then take the product of the two as the BRS (brokerage report sentiment), as shown the following equation:

$$\text{BRS}_{k,t} = \text{RRA}_{k,t} * \text{RRRS}_{k,t}. \quad (3)$$

The difference is that attention in this paper is calculated daily (in literature [4], it is calculated monthly). Specifically,

The first step is to calculate the preliminary sentiment index: Sentiment^1 . Specifically, first standardize the market state indicators, then perform principal component factor analysis on the indicators, and select the three principal components with the highest variance explanation as weights; finally, the factor load, that is, the coefficient of Sentiment^1 , is obtained after weighted average.

the attention index $\text{RRA}_{k,t}$ is constructed by the ratio of the absolute number of reports of stock k on that day to the total number of all reports of stock A-share market on that day. The calculation method is shown as follows:

$$\text{RRA}_{k,t} = \frac{N_{k,t}}{N_{A,t}}, \quad (4)$$

where $N_{k,t}$ is the total number of reports of stock k on day t and $N_{A,t}$ is the total number of reports of all stocks in the A-share market on day t .

Stock k report rating sentiment ($\text{RRRS}_{k,t}$) comprehensively considers the two factors of base rating and rating change. Specifically, the assignment of base rating and rating change is shown in Table 1, and the calculation method is as follows:

$$\text{RRRS}_{k,t} = \frac{\sum_1^{N_{k,t}} R_{k,t} \cdot C_{k,t}}{N_{k,t}}, \quad (5)$$

where $R_{k,t}$ is the base rating and $C_{k,t}$ is the rating change. When there are multiple ratings for an individual stock in a single day, the research report ratings are averaged.

TABLE 1: Discrete assignment of base ratings and rating changes.

Variable	Assignment	Variable	Assignment		
Base rating	Strong buy	Rating Change	Up		
	Buy			1.2	
	Long-term buy			1.7	
	Overweight			1.6	
	Buy with caution		1.5	First	1.1
	Highly recommended		1.4		
	Recommend		1.3	Maintenance	1.0
	Cautious overweight		1.2		
	Hold		1.1		
	Neutral		1.0		
	Recommended with caution		0.9	Down	0.9
	Wait and see		0.8		
	Underweight		0.7		
	Sell		0.6		

3.1.3. *News Sentiment.* News on the Web is a long text, and headlines alone cannot accurately and completely express the text. Therefore, we first generate a summary of the news, obtain the accurate intention of the text, and then calculate the NS (news sentiment).

(1) *News Summary Generation.* We use the architecture of Seq2Seq [25] to generate the summary, where the encoder takes a sequence as input, encodes the information in the sequence as a semantic vector, and then outputs the summary text through the decoder. The model is shown in Figure 2.

The encoder is bidirectional long short-term memory (Bi-LSTM) network. The input news is represented as $x = \{x_1, x_2, x_3, \dots, x_N\}$, and we encode x into hidden state vectors with

$$h = \text{BiLSTM}(x). \quad (6)$$

$h = \{h_1, h_2, h_3, \dots, h_N\}$. Specially, h_N is the result of the merger of two-way last hidden states. The decoder part uses LSTM, whose initial state s_0 is the output h_N of the encoder. On the step t , the decoder receives the previous decoder state s_{t-1} and the previously generated token y_{t-1} , and the decoder current state s_t is calculated as follows:

$$s_t = \text{LSTM}(s_{t-1}, y_{t-1}). \quad (7)$$

This method only uses s_t to connect the encoder and decoder, so the encoder needs to compress the entire sequence information into a fixed-length vector, which is limited. As the length of the input sequence increases, the information entered first is diluted by the information entered later. For better decoding, we use the attention mechanism [26] to instruct the decoder to generate the next word through the probability distribution on the source word. The attention distribution a^t can be calculated by s_t and h_i :

$$\begin{aligned} e_i^t &= v^T \tanh(W_h h_i + W_s s_t), \\ a^t &= \text{soft max}(e^t), \end{aligned} \quad (8)$$

where v, W_h, W_s are learnable parameters and a^t computes the context vector h_t^* :

$$h_t^* = \sum_i a_i^t h_i. \quad (9)$$

h_t^* contains decoding information, and we finally get the probability distribution P_{vocab} of the output words through h_t^* :

$$P_{\text{vocab}}(y_t | x, y_{<t}) = \text{soft max}(V' \tanh(V[s_t; h_t^*])), \quad (10)$$

where V' and V are learnable parameters.

(2) *News Sentiment Computing Based on Rules.* Referring to the study of Qi [27], we construct relevant semantic rules to dig out the real emotions of semantic words in different contexts. Specifically, according to the number of semantic words, the text is divided into multiple clusters, and the emotional value of the text is the sum of the semantic values of the clusters. The calculation formula is as follows:

$$U_a = \prod L_s \times S, \quad (11)$$

$$U_b = \left(\prod L_n \times (0.1 \times C_n - 1)^{c_n} + \prod L_s \right) \times S, \quad (12)$$

where U_a is the sentiment value of the cluster without negative words; U_b is the sentiment value of the cluster with negative words; S is the semantic value of the emotional word; L_s is the degree value of the degree adverb that modifies semantic words; L_n is the degree value of the degree adverb that modifies the negative word; C_n is the number of negative words. Then, the average of all the news semantic of the day was calculated, namely, news semantic (NS).

$$\text{NS} = \sum_1^n U_a + \sum_1^n U_b. \quad (13)$$

3.1.4. *We-Media Sentiment.* The data we chose to calculate WMS (We-Media sentiment) came from the stock BBS (Bulletin Board System). Retail investors communicate in

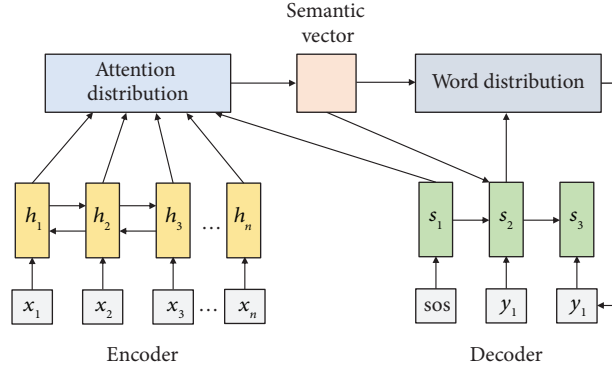


FIGURE 2: Abstract generation of a news text based on Seq2Seq.

the form of posts and replies, and the information they publish is usually short. After analysis, we believe that a post, including the post information and the reply information, represents the investment sentiment, so we combine the short text of the post and reply into a long text and calculate the sentiment of the long text. Refer to formulas (11)–(13) for WMS calculation.

3.1.5. Multisource Information Fusion. We use a multilayer perceptron [23] to fuse the output of four different emotion vectors and then use the aggregate vector as the first input of the stock prediction LSTM. ISV calculation is as follows:

$$\text{ISV} = v^T \sigma(W \cdot [MS, \text{BRS}, NS, \text{WMS}] + b). \quad (14)$$

3.2. Stock Price Prediction Model Based on LSTM. LSTM takes the output of the encoder as the input of $t = -1$ time step to guide the prediction of subsequent stock price, and the output m_{t-1} of LSTM at $t - 1$ time step is the input of t time step. In the training phase, for ISV and stock price $S = (S_0, \dots, S_N)$, the probability formula for predicting the stock price of the next trading day is as follows:

$$\begin{aligned} x_{-1} &= \text{ISV}, \\ x_t &= S_t, \quad t \in \{0, \dots, N - 1\}, \\ p_{t+1} &= \text{LSTM}(x_t), \quad t \in \{0, \dots, N - 1\}, \end{aligned} \quad (15)$$

where x_{-1} and x_t , respectively, represent the inputs of LSTM in time steps -1 and t , S_0 represents the closing price of stocks at the beginning trading day, and S_N represents the closing price of stocks in the end trading day. The loss function formula of the whole model is as follows:

$$L(I, S) = - \sum_{t=1}^N \log p_t(s_t), \quad (16)$$

where I represents ISV.

4. Experiment

4.1. Data and Preprocessing. The experimental data were selected from January 1, 2020, to December 31, 2020, excluding new shares and long-term suspended stocks. All the web texts are captured by scrapy crawler framework and preprocessed by word segmentation and removal of stop words.

- (1) Macroeconomic data: considering the lag of macroeconomic data, the data is selected from September 2019 to September 2020, and the data source is the WIND database (<https://www.wind.com.cn/NewSite/edb.html>).
- (2) Brokerage research reports: a total of 32724 reports on 2365 A-share companies released by 63 securities institutions are included. Data on the number of published reports, date of release, title, basic rating, and change of rating were obtained from East-money (<https://www.eastmoney.com/>). The brokerage rating confidence data are shown in Table 2.
- (3) News: we selected the news on the four authoritative websites of China Securities Network (<http://www.cs.com.cn/>), Sina Finance (<https://finance.sina.com.cn/>), Netease Finance (<https://money.163.com/>), and Securities Times (<http://www.stcn.com/>) as news data sources and captured the content including news headlines, release time, and news content. After sorting, a total of 96,532 news articles were obtained.
- (4) We-Media: the We-Media data came from Guba (<https://guba.eastmoney.com/>) and Xueqiu (<https://xueqiu.com/>), two BBS where Chinese retail investors discuss stocks. The number of “We-Media” texts after splicing is 183,938.

4.2. Baseline. In this paper, stock returns are used as the research object to predict whether the returns obtained by individual stocks in a certain period in the future are positive, negative, or flat. To ensure robustness, all data are

TABLE 2: Descriptive statistical results of brokerage rating sentiment.

Sentiment	Optimistic (RRRS > 1)	Neutral (RRRS = 1)	Pessimistic (RRRS < 1)
Number	31725	107	892
Proportion (%)	96.948	0.327	2.725

standardized according to the returns of the market in this period. To measure the advantages and disadvantages of the model from different perspectives, this paper selected SVM, LSTM model, RrmsNet [4], and SenticNet [5] as the benchmark methods for comparison with our work.

4.3. *Metrics.* In the experiment, accuracy and F1-score are adopted to evaluate the performances of each method. Let n denote the total number of samples and N_c denote the number of samples whose true label is c . These metrics are defined as follows:

$$\text{accuracy} = \sum_{c \in y} \frac{n_{cc}}{N}, P_c = \frac{n_{cc}}{n_{cc} + \sum_{p \neq c} n_{pc}}, R_c = \frac{n_{cc}}{n_{cc} + \sum_{p \neq c} n_{cp}}, F_1 = \sum_{c \in y} \frac{N_c}{N} \frac{2P_c R_c}{P_c + R_c}, \quad (17)$$

where $y = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$, N_c is the total number of samples in class c , N is the total number of samples, and n_{ij} is the number of samples whose true label is i and the predicted label is j . P_c and R_c are precision and recall.

5. Results

5.1. *Main Result.* Due to the timeliness of information dissemination, we choose the 5th, 15th, and 30th as the window period for experimental observations. Table 3 shows the detailed results of the comparative experiment. In general, the method we proposed has achieved the best results in both accuracy and F1-score, which shows that investor sentiment vectors combined with multisource information can effectively improve the performance of stock price prediction.

Figure 3 shows the accuracy and F1-score for different time windows. It can be seen from the figure that over time, the accuracy and F1-score of all methods have declined. Taking our method as an example, the accuracy rates on the 5th, 15th, and 30th days are 0.749, 0.693, and 0.668, respectively, and the F1-score are 0.723, 0.699, and 0.641, respectively. There are two reasons for this. First, all methods, no matter whether third-party information is included, are based on historical stock prices to predict future stock prices. As a result, the greater the time, the greater the uncertainty of the prediction. Second, investor sentiment based on the comprehensive calculation of different information is essentially an expression of information dissemination, so its influence on stock price prediction will weaken over time. This is consistent with the theory of information communication, that is, the longer the time, the weaker the influence of information.

5.2. *Ablation Experiments.* To better observe the influence of MSI, BRS, NS, and WMS on the performance of stock price prediction, an ablation experiment was carried out in this paper. The main idea is to remove one of the above indicators, respectively, to obtain four models Without_MSI, Without_BRS, Without_NS, and Without_WMS. Then, the accurate value and F1-score are compared with the Our_full

model. The larger the difference, the greater the influence and contribution. The ablation results are shown in Table 4. In general, excluding any index, the accuracy and F1-score are lower than the Our_full model, which indicates that the four indicators measuring investor sentiment have a positive impact on the stock price prediction. Among them, the Without_BRS model excluding the BRS indicator has the largest gap compared with the Our_full model, which shows that, among the four indicators, BRS has the greatest impact on the stock price prediction. There are two reasons. First, BRS, as a professional brokerage research report, is more easily recognized by shareholders. Second, compared with other sentiment indicators, brokerage research reports will directly give buy or sell recommendations, more direct.

Next, we remove the two indicators for further testing the performance of the model. Specifically, one is to get the Without_MSI_BRS model by removing MS and BRS simultaneously and the other is to get the Without_NS_WMS model by removing NS and WMS simultaneously. The reason for this is that MS and BRS information is more formal and comes from official or institutional sources, while NS and WMS information comes from news and comments on the Internet, which is more casual and free. The results show that the gap between the Without_MSI_BRS and Our_full model is larger, which shows that although the amount of news and commentary information from the Internet is greater, the price of stocks is more affected by official economic indexes and brokerages.

Finally, we compare the influence of different fusion methods of four indicators on the stock price prediction. Concatenation model splices four indicators into one-dimensional vectors as LSTM inputs, and the results show that its performance is inferior to Our_full model, indicating that the fusion method proposed in this paper is more suitable for stock price prediction tasks.

5.3. *Long-Term Stock Price Impact Analysis.* To examine whether there is a long-term effect on the impact of information on stock prices, we choose to conduct experiments on the 45th, 60th, and 90th window periods. The

TABLE 3: Stock returns forecast results.

Method	5 days		15 days		30 days	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
SVM	0.672	0.633	0.623	0.607	0.583	0.567
LSTM	0.710	0.684	0.634	0.618	0.599	0.578
RrmsNet [4]	0.727	0.718	0.656	0.643	0.618	0.604
SenticNet [5]	0.719	0.704	0.666	0.640	0.611	0.597
Ours	0.749	0.732	0.693	0.699	0.668	0.641

Bold values represent the best performance.

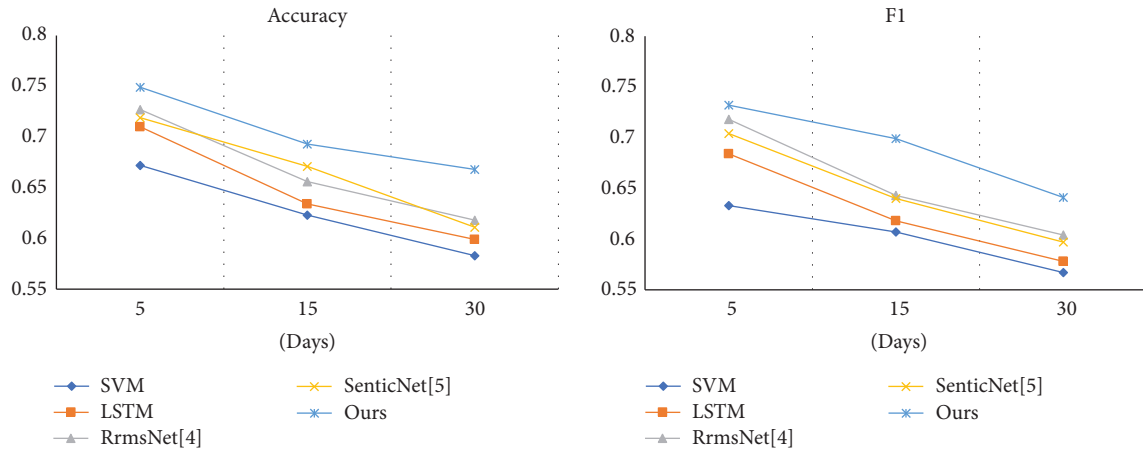


FIGURE 3: Comparison of results in different time windows.

TABLE 4: Results of ablation experiments.

Method	5 days		15 days		30 days	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
Without_MSI	0.732	0.719	0.681	0.671	0.639	0.622
Without_BRS	0.688	0.713	0.646	0.670	0.610	0.591
Without_NS	0.721	0.719	0.685	0.665	0.638	0.624
Without_WMS	0.730	0.725	0.680	0.667	0.638	0.617
Without_MSI_BRS	0.701	0.698	0.640	0.642	0.630	0.609
Without_NS_WMS	0.732	0.723	0.677	0.668	0.633	0.621
Concatenation	0.736	0.725	0.687	0.672	0.642	0.626
Our_full	0.749	0.732	0.693	0.699	0.668	0.641

Bold values represent the best performance.

results are shown in Table 5. It can be seen from the table that the accuracy and F1-score of all methods are between 0.50 and 0.60, and it does not clearly show which method has better performance. Through the analysis of the case, it is found that, in the 45th, 60th, and 90th day time window after the information is released, the stock price performance of the predicted stocks is unstable and even presents a certain degree of randomness.

First, the basic methods of stock price prediction in other models except SVM are all based on the LSTM model, and the LSTM model itself has the problem that the input sequence is too long and the gradient disappears. Further, the investor sentiment vector aggregated by our proposed method is used as the initial input of the LSTM model ($t = 1$). As the observation window expands to 45 days or even 90 days, the influence of investor sentiment vectors on subsequent time steps gradually weakens.

Second, in the self-media era, the update cycle of market-related information is relatively short. Among the four information sources we selected, the MSI cycle is updated monthly, the BRS update cycle is about 20 days, the NS update cycle is about 2 weeks, and the WMS is updated daily, as shown in Table 6. In other words, the longest period of all information is 30 days, which means that new news will overwrite old news and affect investors' decision-making.

Finally, China's A-share market is a semiclosed and immature market. Investors' decision-making is often affected by the latest information, leading to frequent transactions and short holding periods. According to statistics, the average holding period of individual investor accounts in the A-share market is less than 20 trading days; even for investment institutions, the average holding period is about 30–40 trading days. The characteristics of the market determine the direction of the market.

TABLE 5: Stock return forecast results in the long term.

Method	45 days		60 days		90 days	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
SVM	0.587	0.533	0.567	0.564	0.544	0.512
LSTM	0.516	0.541	0.526	0.562	0.507	0.531
RrmsNet [4]	0.542	0.550	0.533	0.544	0.512	0.507
SenticNet [5]	0.568	0.549	0.589	0.551	0.527	0.514
Ours	0.548	0.539	0.542	0.577	0.525	0.530

Bold values represent the best performance.

TABLE 6: Stock return forecast results in the long term.

	MSI	BRS	NS	WMS
Update cycle	A month	Three weeks	Twelve days	One day

In summary, the investor sentiment vector calculated from the day of information collection has a limited impact on the stock price 45 days or even 90 days later, which is also the reason for the poor performance of the model.

6. Conclusion

The relationship between investor sentiment and the stock price has always been a hot research topic. In the era of big data, the channels for investors to obtain information have changed from research reports and news dominated by securities brokers to We-Media information. Multiple sources of information have brought new changes to measures of investor sentiment. Based on multisource information fusion, this paper proposes a new measurement method of investor sentiment and incorporates the new investor sentiment into the framework of stock price prediction. In the experiment with the data of China A-shares from January 2020 to December 2020, the results show that (1) investor sentiment is an important factor affecting stock price fluctuations, (2) among the different indicators of investor sentiment, brokerage report sentiment has the greatest impact on stock prices, and (3) multilayer perceptrons can better integrate emotional indicators.

Data Availability

All data in this paper come from public information on the Internet.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Natural Science Foundation of Jiangxi Province (grant no. 20212BAB202016) and the Science and Technology Research Project of Jiangxi Provincial Department of Education (grant no. GJJ200318).


References

- [1] D. B. Dai, Y. S. Lan, T. J. Fan, and M. Zhao, "Stock Forecast with Investors Sentiment by Text Mining and Machine Learning," *China Soft Science*, vol. 4, pp. 166–175, 2019.
- [2] G. A. Akerlof and R. J. Shiller, *Animal Spirits: How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*, Princeton university press, Princeton, NJ, USA, 2010.
- [3] J. You and J. Wu, "Spiral of Silence: Media Sentiment and the Asset Mispricing," *Economic Research Journal*, vol. 47, 2020.
- [4] H. Song, D. Peng, and X. Huang, "Incorporating Research Reports and Market Sentiment for Stock Excess Return Prediction: A Case of Mainland China," *Scientific Programming*, vol. 2020, Article ID 8894757, 7 pages, 2020.
- [5] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: a case of Hong Kong," *Information Processing & Management*, vol. 57, no. 5, Article ID 102212, 2020.
- [6] C. Polk and P. Sapienza, "The stock market and corporate investment: a test of catering theory," *Review of Financial Studies*, vol. 22, no. 1, pp. 187–217, 2008.
- [7] B. Debata, S. R. Dash, and J. Mahakud, "Monetary Policy and Liquidity: Does Investor Sentiment matter?" *IIMB Management Review*, vol. 33, 2021.
- [8] K. L. Fisher and M. Statman, "Investor sentiment and stock returns," *Financial Analysts Journal*, vol. 56, no. 2, pp. 16–23, 2000.
- [9] Z. Da, J. Engelberg, and P. Gao, "The sum of all FEARS investor sentiment and asset prices," *Review of Financial Studies*, vol. 28, no. 1, pp. 1–32, 2015.
- [10] X. J. Meng, X. L. Meng, and Y. Y. Hu, "Research on investor sentiment index based on text mining and Baidu index," *Macroeconomics*, vol. 38, no. 1, pp. 144–153, 2016.
- [11] H. Y. Yin and X. Y. Wu, "Predictive effect of high-frequency investor sentiment on the intraday stocks return," *China Industrial Economics*, vol. 8, no. 8, pp. 80–98, 2019.
- [12] H. Song, D. Peng, X. Huang, and J. Feng, "Research on weibo hotspot finding based on self-adaptive incremental clustering," *Journal of Shanghai Jiaotong University*, vol. 24, no. 3, pp. 364–371, 2019.

- [13] N. Oliveira, P. Cortez, and N. Areal, "The impact of micro-blogging data for stock market prediction: u," *Expert Systems with Applications*, vol. 73, pp. 125–144, 2017.
- [14] C. Zhao, Y. W. Ye, and M. H. Yao, "Stock volatility forecast based on financial text emotion," *Computer Science*, vol. 47, no. 5, pp. 87–91, 2020.
- [15] Z. G. Yi and N. Mao, "Research on the measurement of investor sentiment in Chinese stock market: the CICSII's construction," *Journal of Financial Research*, vol. 11, pp. 174–184, 2009.
- [16] R. W. Ma and N. Zhang, "The construction of investor sentiment index for China's stock market: based on the panel data of shanghai A share companies," *Journal of Central University of Finance & Economics*, vol. 7, pp. 42–49, 2015.
- [17] N. Pröllochs, S. Feuerriegel, and D. Neumann, "Negation scope detection in sentiment analysis: decision support for news-driven trading," *Decision Support Systems*, vol. 88, pp. 67–75, 2016.
- [18] T. O. Sprenger, A. Tumasjan, P. G. Sandner, and I. M. Welp, "Tweets and trades: the information content of stock microblogs," *European Financial Management*, vol. 20, no. 5, pp. 926–957, 2014.
- [19] H. K. Sul, A. R. Dennis, and L. I. Yuan, "Trading on twitter: using social media sentiment to predict stock returns," *Decision Sciences*, vol. 48, no. 3, pp. 454–488, 2017.
- [20] M. S. Checkley, D. A. Higón, and H. Alles, "The hasty wisdom of the mob: how market sentiment predicts stock market behavior," *Expert Systems with Applications*, vol. 77, pp. 256–263, 2017.
- [21] W. Antweiler and M. Z. Frank, "Is all that talk just noise? The information content of Internet stock message boards," *The Journal of Finance*, vol. 59, no. 3, pp. 1259–1294, 2004.
- [22] S. R. Das and M. Y. Chen, "Yahoo! For amazon: sentiment extraction from small talk on the web," *Management Science*, vol. 53, no. 9, pp. 1375–1388, 2007.
- [23] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and translate," 2014, <https://arxiv.org/abs/1409.0473>.
- [24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [25] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, Salt Lake City, UT, USA, June 2018.
- [26] A. M. Rush, S. Chopra, and J. Weston, "A neural attention model for abstractive sentence summarization," 2015, <https://arxiv.org/abs/1509.00685>.
- [27] T. F. Qi and H. X. Jiang, "Exploring stock price trend using Seq2Seq based automatic text summarization sentiment mining," *Management Review*, vol. 33, no. 5, p. 257, 2021.

Research Article

Regional Atmospheric Light Optimisation Algorithm for Heterogeneous Image Dehazing

Haoqiang Wu,^{1,2} Yiran Fu,^{1,2} Quanxing Zha,^{1,2} Aidong Chen,^{1,2,3} and Hongyuan Jing^{1,2,3} 

¹Multi-Agent System Research Centre, Beijing Union University, Beijing 100101, China

²College of Robotics, Beijing Union University, No. 4 Gongti North Road, Chaoyang District, Beijing 10002, China

³Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China

Correspondence should be addressed to Hongyuan Jing; jqrhongyuan@buu.edu.cn

Received 23 July 2021; Revised 25 August 2021; Accepted 1 November 2021; Published 19 November 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Haoqiang Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Under foggy and other severe weather conditions, image acquisition equipment is not effective. It often produces an image with low contrast and low scene brightness, which is difficult to use in other image-based applications. The dark channel prior dehazing algorithm will cause the brightness of the image to decrease and sometimes introduce halos in the sky area. To solve this problem, we proposed a region similarity optimisation algorithm based on a dark channel prior. First, a vector comprising RGB layer dark channel value was obtained as the original atmospheric ambient light, and then, the proposed regional similarity linear function was used to adjust the atmospheric ambient light matrix. Next, the transmittance of different colour channels was derived and the multichannel soft matting algorithm was employed to produce more effective transmittance. Finally, the atmospheric ambient light and transmittance were substituted into the atmospheric scattering model to calculate clean images. Experimental results show that the proposed algorithm outperformed the existing mainstream dehazing algorithms in terms of both visual judgement and quality analysis with nonhomogeneous haze datasets. The algorithm not only improves the image details but also improves the brightness and saturation of the dehazing result; therefore, the proposed algorithm is effective in the restoration of the hazy image.

1. Introduction

Fog is a near-surface atmospheric weather phenomenon caused by the desublimation of suspended water droplets in the air. Foggy conditions cause considerable inconvenience to human life and production, especially in imaging. Because of the light scattering caused by various particles in a foggy atmosphere, such as PM_{2.5}, image acquisition equipment tends to produce images with fewer details, low contrast, low scene brightness, and whitish hue, which not only renders a poor intuitive visual experience but also affects the normal operation of the work system in image-related fields, such as image recognition. Thus, research on haze removal technology is valuable and significant.

At present, image dehazing technology is mainly divided into three categories [1–3]: image enhancement methods, physical model-based methods, and deep learning-based methods using neural networks. Among them,

representative image enhancement algorithms [4–9] include histogram equalisation, homomorphic filtering, wavelet transform, and MSR algorithm based on the Retinex (SSR) theory.

The histogram equalisation algorithm nonlinearly stretches and rearranges the dense part of the whole grey histogram to achieve approximately uniform distribution, thus enhancing the local image contrast. The Retinex model proposed by Edwin Land was based on the colour constancy theory [4]. Based on the theory of the single-scale SSR algorithm, Ma and Wen [5] combined edge detection within a Gaussian filtering algorithm to amend the estimated value of the reflection component and mitigate the loss of image edge information. Zheng et al. [6] proposed an adaptive structure decomposition-integrated multiexposure image fusion dehazing method. It linearly adjusts the underexposed images through gamma correction to enhance the visual effect after dehazing, but the calculation complexity is high.

Zhu et al. [7] proposed an image fusion-based algorithm that reduces haze effects via gamma correction. This algorithm can enhance the saturation of images and can be applied to geographic remote sensing and underwater images. Shu et al. [8] proposed a hybrid regularised variational framework to improve the scene depth of images along with dehazing. In addition, a two-step correction mechanism is employed to address the problem of colour distortion in the sky area.

Physical model-based methods use the traditional atmospheric scattering model [10] as the main research object and then estimate atmospheric ambient light and transmittance via a priori hypothesis to make them similar to the real scene, thus realising image defogging processing [11–24]. Kaiming He et al. [13] proposed a single-image dehazing algorithm using the dark channel prior theory. This method adopts soft matting method to refine the transmittance and has a remarkable effect. Later, He et al. [14] substituted guided filtering for soft matting in transmittance refinement to improve the operating efficiency, but phenomena such as overall bluish and caliginous hue and colour distortion in the sky area still exist. Zhu et al. [15] obtained the mapping relation by training a linear model built from a colour attenuation prior. This algorithm achieves less colour distortion and outperforms most traditional algorithms. Meng et al. [16] employed a regularisation algorithm to effectively constrain the boundary of the sky region to obtain a restored image. This method causes image detail loss and artifacts in the image boundary. Sulami et al. [17] assumed that the scene chrominance was uncorrelated with the transmittance and calculated the transmittance using statistical laws, resulting in a poor effect in images with relatively single colour features. Berman et al. [18] proposed a nonlocal prior within each type of pixel cluster; the pixel points will have a linear relationship according to the difference in fog concentration. They estimated the scene depth based on this prior and accordingly restored the fog-free image. Ju et al. [19] introduced a novel light absorption coefficient parameter to attain an enhanced ASM (EASM). This model increases the visibility of hazy images while dehazing outdoor hazy images. Later, Ju et al. [20] proposed an image dehazing method using blended priors. This algorithm combines multiple modules, including the atmospheric light estimation module to increase brightness and multiple prior constraint modules to create a nonlocal prior, local prior, and global prior, which improve the brightness and saturation of images. Wang et al. [21] proposed an additional channel method based on quad-tree subdivision and refined the transmittance using Gaussian blurring. These measures prevent the phenomena of oversaturation and halo effects from happening and tackling the problem of distortion in the brightest areas.

In recent years, machine learning has become an important topic for its wide application and methods based on deep learning and neural networks have also made progress in image defogging [25–34]. Chen et al. [25] proposed a single-image dehazing method based on CNN, obtained the detailed features of foggy images via convolution layer operation and multiscale convolution kernel operation, and then fitted the transmittance map using nonlinear

regression. In contrast to CNN, Cai et al. [26] proposed a trainable end-to-end transmission medium estimation model called DehazeNet, which directly learns and estimates the mapping relationship between foggy images and transmittance. However, this algorithm still uses atmospheric light as a global constant, leading to a loss in the precision of the image to a certain degree. Li et al. [27] improved the traditional atmospheric scattering model and used a convolutional neural network to estimate the residual image containing atmospheric light and transmittance. Qin et al. [28] proposed an end-to-end feature fusion dehazing network. This network structure can adaptively learn different weights of different-level feature information. The results show significant dominance in image detail restoration and colour fidelity. Liu et al. [29] designed and enhanced the CNN learning method for ship detection, which incorporated nonmaximum suppression to accurately obtain the results and construct a mixed loss function balance class during the training phase. Like the relative fine dehazing results as those aforementioned networks have, the deep learning-related algorithm relies considerably on big data; this is difficult to implement under sparse samples.

Algorithms, such as the dark channel a priori based on the traditional atmospheric scattering model, ignore most of the image detail information when estimating the atmospheric ambient light. To solve this problem, we made the following modifications based on the dark channel a priori theory: (1) The dark channel values were obtained for each of the three RGB colour channels of the hazy image, and the regional mean value of each pixel in each dark channel was considered as the corresponding atmospheric ambient light value, after which the entire atmospheric ambient light was linearly transformed with its top 1% value to obtain the final atmospheric ambient light value. This reduces the influence of image noise to a certain extent. (2) In this study, we proposed to refine the transmittance through guided filtering and selected the V channel in the HSV with the best effect as the guided map for transmittance optimisation through experiments to obtain a transmittance closer to the real scene. In this study, we ran experiments on the O-Haze dataset, NTIRE 2021 dataset, and real images and analysed the results from both subjective and objective aspects. The results show that the proposed algorithm effectively improves the brightness and saturation of the image and the haze-free image after dehazing has high colour contrast and good visual effect and also improves the problem of blue and dark images after the dark channel a priori algorithm processing.

2. Background

Under foggy and other severe weather conditions, the light will scatter when it meets suspended small particles in the air during propagation, resulting in degraded images produced by image acquisition devices. Nayer and Narasimhan improved the atmospheric scattering model [10] in 1999, and this model (Figure 1) is widely used in computer vision and computer graphics.

It can be formulated as

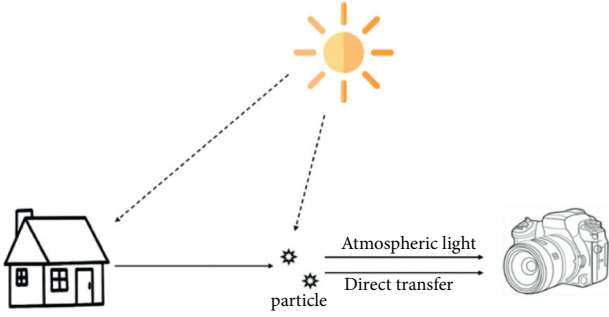


FIGURE 1: Atmospheric scattering model.

$$I(X) = J(X)t(X) + A(1 - t(X)), \quad (1)$$

where $I(X)$ refers to the hazy image received by imaging equipment, $J(X)$ is the real-scene radiation, A is the atmospheric ambient light value, and $t(X)$ is the transmittance, which can be expressed as

$$t(X) = e^{-\beta d(X)}, \quad 0 < t(X) \leq 1, \quad (2)$$

where β is the atmospheric scattering coefficient and $d(X)$ is the distance between the actual scene and the photographed point, that is, the depth of the field.

As per the formula, $J(X)$ can be obtained via formation simplification after estimating atmospheric ambient light A and transmittance $t(X)$ of input $I(X)$.

Kaiming He et al. [13] conducted experiments on large quantities of outdoor haze-free images and found that in the vast majority of nonsky regions in haze-free images, some pixels in a certain region have at least one colour channel with a rather low value. In dark channels, they pointed out that dark channel values tend to zero in nonsky areas. The dark channel value in a local patch $\Omega(X)$, with the centre pixel x , is denoted by $J^{\text{dark}}(X)$ as follows:

$$J^{\text{dark}}(x) = \min_{c \in \{r, g, b\}} \left(\min_{y \in \Omega(x)} (J^c(y)) \right) \rightarrow 0. \quad (3)$$

In practice, small particles inevitably exist in the atmosphere. The presence of haze is fundamental for observers to perceive scene depth. Therefore, after minimisation on both sides of equation (3), the parameter ω ($0 < \omega \leq 1$) is introduced to maintain a certain amount of haze, thus enhancing the visual experience of observers, as shown in

$$\tilde{t}(X) = 1 - \omega \min_c \left(\min_{y \in \Omega(X)} \left(\frac{I^c(Y)}{A^c} \right) \right), \quad (4)$$

where c represents the RGB colour channels, $c \in \{r, g, b\}$, I^c is the hazy image of each colour channel layer, and A^c is the atmospheric ambient light of each colour channel layer.

The dark channel prior is subject to limitations in that it does not hold in the sky area and white object area; therefore, results with images containing such areas will display local image distortion and a widespread halo in the sky area.

3. Proposed Algorithm

Due to the unrobustness of the dark channel prior, its dehazing results are subject to problems such as bluish hue and colour distortion in the sky area. The proposed algorithm refines the estimation of the atmospheric ambient light and transmittance based on the original dark channel prior algorithm through the region similarity correction, thus obtaining more effective estimate values, which significantly improves the dehazing effect. The flowchart of the proposed algorithm is shown in Figure 2.

As shown in Figure 2, (a) is the input hazy image, we first select the low brightness pixels in a moving windows from (a) as the dark channel image (b). Then, we use the region similarity correction method to estimate the enhanced atmospheric light image (c). By using the dark channel prior, the rough transmittance map can be calculated as (d). In addition, the input image are transferred to HSV domain (e) to get the intensify map (f). The detailed transmittance map (g) is generated by a guide filter which employed the intensity map (f) as the guide map. Finally, the enhanced atmospheric light (e) and detailed transmittance (g) are submitted to the atmospheric scattering model to get the dehazing output (h).

3.1. Estimating the Dark Channel Value. For a hazy image with a size of $m \times n$, we designed a window $\Omega(X)$ centered around it and chose the average value of the lowest grey value for each pixel in the R , G , and B colour channels, respectively. Furthermore, within the range of $\Omega(X)$, because of the dark channel value of this pixel, we obtain an $m \times n \times 3$ RGB dark channel matrix $J^{c,\text{dark}}(X)$ (equation (6)):

$$J^{c,\text{dark}}(X) = \min_{Y \in \Omega(X)} (J^c(Y)), \quad (5)$$

$$c = R, G, B,$$

$$J^{\text{dark}}(X) = \{J^{c,\text{dark}}(X)\}. \quad (6)$$

The size of $\Omega(X)$ has an impact on the dark channel image; the larger its value, the darker the pixel points it contains and the darker the generated image; the experimental effect will not be suitable. The smaller the value, the less the information of the dark channel included, and the difference with the original image is not significant. Therefore, $s = 31$ was chosen in this study to achieve the best balance.

We further compute the atmospheric ambient light using it.

3.2. Estimating the Atmospheric Ambient Light. To better realise region selection, for each pixel in the obtained $J^{c,\text{dark}}(X)$ as the starting point, we determined a square window $\Omega(X)$ from four different directions, including the upper left, upper right, lower left, and lower right of the pixel independently to calculate the mean dark channel value from within as the atmospheric light $A^c(X)$ of this pixel and

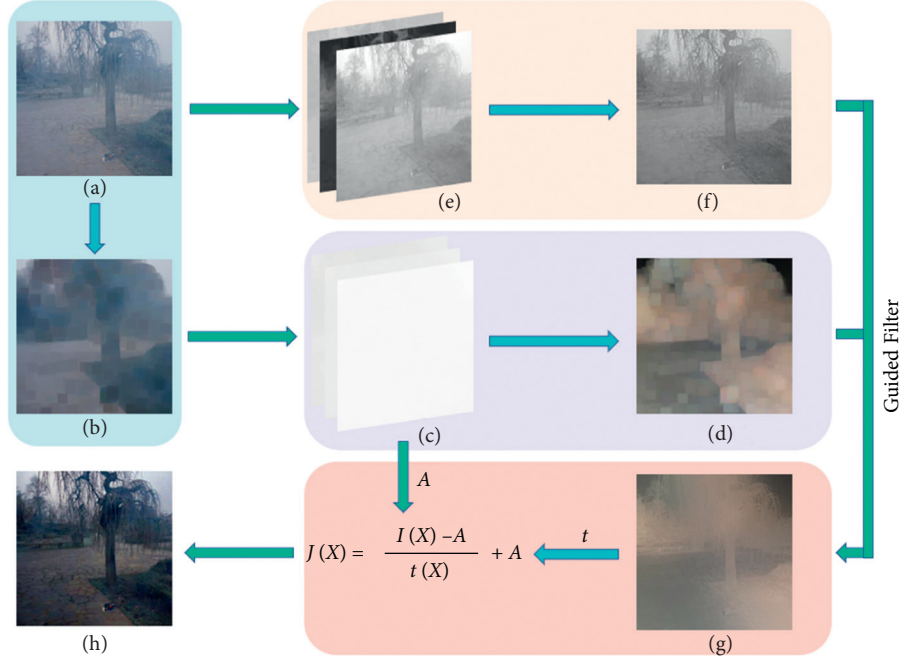


FIGURE 2: Flowchart of the proposed method.

experiment with their dehazing effect. The results show that the lower right direction achieves the best; therefore, we chose this direction as the final method, as shown in

$$A^c(X) = \frac{\sum_{i=x}^{x+s'-1} \sum_{j=y}^{y+s'-1} J^{c,\text{dark}}(i, j)}{s'^2}, \quad (7)$$

where $J^{c,\text{dark}}(i, j)$ represents the dark channel value of the pixel with coordinates (i, j) , $c = R, G, B$, s' is the length of the neighbourhood window, and x and y are the abscissa and ordinate, that is, $\begin{cases} 1 \leq x \leq m \\ 1 \leq y \leq n \end{cases}$. Provided that the size of the square window size at the image edge is less than $\Omega(X)$, then we select the window with side length $\min\{x, y, m - x + 1, n - y + 1\}$.

Larger atmospheric ambient light values will make the image brighter, while too small values will result in a darker image overall and affect the accuracy of the estimated transmittance $t(X)$. This is particularly obvious in the dense haze region. Therefore, in this study, we proposed an improvement scheme, in which the obtained $A^c(X)$ values are first arranged in a descending order and the first 1% of them are considered as the ideal values of the atmospheric ambient light to ensure that the atmospheric ambient light is more consistent with the real scene as shown in

$$N = m \times n \times 0.01, \quad (8)$$

where the positive integer N refers to the corresponding value of the 1% pixel of the whole image (rounded down).

Refactoring the $m \times n$ vector rearranged from $A^c(X)$ as $Re^c(h_0)$, then the pixel with coordinates (x, y) can be formulated as follows:

$$\begin{aligned} h_0 &= (x - 1) \times n + y, \\ Re^c(h_0) &= A^c(x, y), \end{aligned} \quad (9)$$

where h_0 is the subscript position of the A value with the original coordinates (x, y) . Subsequently, we rearrange $Re^c(h_0)$ in descending order and denote the new column vector as $De^c(h)$.

Next, we rearrange the entire $m \times n$ atmospheric light matrix according to their linear proportion. Then, we mark the minimum atmospheric light value as $\min A^c$ and the maximum value of the optimised value $A'^c(X)$ as $\max A^c$. The minimum value $\min A'^c$ of the optimised atmospheric light $A'^c(X)$ is equal to $De^c(N)$. The equation is

$$\frac{\max A^c - A^c(X)}{\max A^c - \min A^c} = \frac{\max A^c - A'^c(X)}{\max A'^c - \min A'^c}. \quad (10)$$

By simplifying (11), we can obtain the rearranged value of each pixel $A'^c(X)$ from the original $A^c(X)$, which leads to

$$A'^c(X) = \max A^c - \frac{(\max A^c - A^c(X))(\max A'^c - \min A'^c)}{\max A^c - \min A^c}. \quad (11)$$

The calculation and rearrangement steps of the atmospheric light are shown in Figure 3.

As shown in Figure 3, a dark channel image (a) was input to obtain the refined atmospheric ambient light image. Considering the pixel point at coordinate (1, 1) as an example, $\Omega(X)$ is selected down to the right (as shown in the red box in (a)) and the mean value of the region $\Omega(X)$ (i.e., (b)) is calculated as the rough atmospheric ambient light value at coordinate (1, 1) (as shown in the red box in the upper left of (c)). Thereafter, we arrange the atmospheric ambient light value of (c) into the right matrix of (d) in

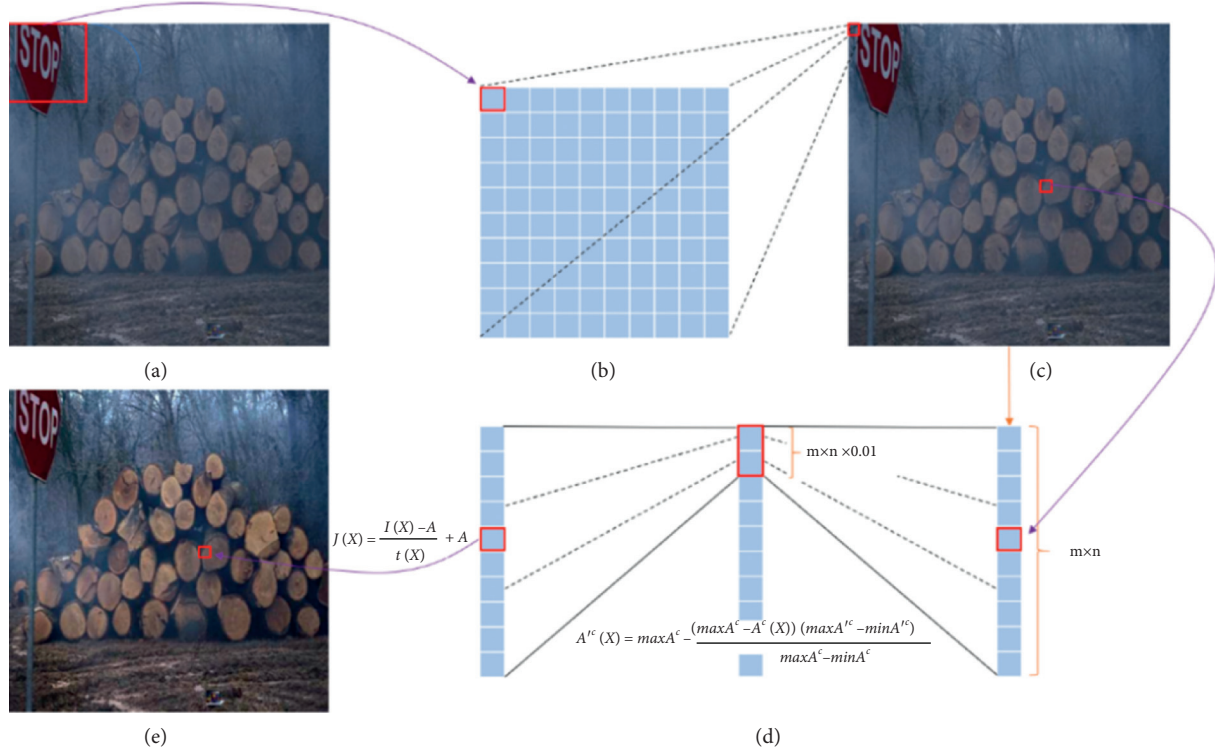


FIGURE 3: Schematic diagram of atmospheric ambient light.

descending order and linearly stretch the first 1% of the pixel points into the new atmospheric ambient light values (as shown in the left matrix of (d)), which is later mapped back to the original coordinate positions for subsequent dehazing (as shown in the change process of the red pixel points from (c) to (e)).

3.3. Estimating the Transmittance. Based on the principle that visible light of different wavelengths has different transmittances, transmittances of different colour channels are obtained in this study, as shown in

$$\tilde{t}^c(X) = 1 - \omega \min_{Y \in \Omega(X)} \left(\frac{I^c(Y)}{A^{I^c}(X)} \right), \quad (12)$$

where $\tilde{t}^c(x)$ represents the three-layer colour channel transmittance matrix.

Experimental analysis indicates that V channel overperforms other choices as the guiding map to employ guided filtering to transmittance \tilde{t}^c , and we obtain a more effective transmittance \tilde{t}^{I^c} , as the following formulas show:

$$\tilde{t}^{I^c} = a_k I_i^v + b_k, \quad \forall i \in w_k, \quad (13)$$

$$a_k = \frac{\text{cov}_k(I^v, \tilde{t}^c)}{\sigma_k^2 + \varepsilon}, \quad (14)$$

$$b_k = \overline{\tilde{t}^c} - a_k u_k, \quad (15)$$

where I^v is the guiding image, that is, the V channel of the hazy image represented in the HSV image system, cov_k is the

covariance of the window w_k , and σ_k^2 and u_k are the variance and mean value of I^v in window w_k , respectively.

The parameter ω also plays a decisive role in the transmittance; the smaller its value, the less obvious the defogging effect. After extensive experiments, we fixed ω to 0.9 and the minimum transmittance was set as 0.1 to avoid excessive enhancement of the image after fog removal.

A comparison of the transmittance \tilde{t}^{I^c} with and without the guided filter is shown in Figure 4.

Thereafter, we submit the improved atmospheric ambient light A^I and transmittance \tilde{t}^{I^d} into equation (1) to obtain clean images.

4. Experimental Results

To evaluate the feasibility and effectiveness of our algorithm more intuitively, we performed comprehensive experiments on a large dataset containing the public O-Haze image dataset, NTIRE 2021 dataset, and real-world images. The O-Haze dataset contains 45 outdoor scene images with the same visual content recorded under fog-free and foggy conditions, from which we stochastically selected six images as the testing set for neural network algorithms and used the others as the training set. The real-world image dataset comprises approximately 150 images, including a public dataset that is available online and self-collected images. The NTIRE 2021 dataset includes 35 pairs of real and outdoor nonhomogeneous hazy images. All operations were performed under the Windows 10 operating system, MATLAB R2020a operating environment, and Intel(R) Core (TM) i7-7500U CPU @2.70 GHz 2.90 GHz processor, with 8 GB RAM hardware configuration. We compared our

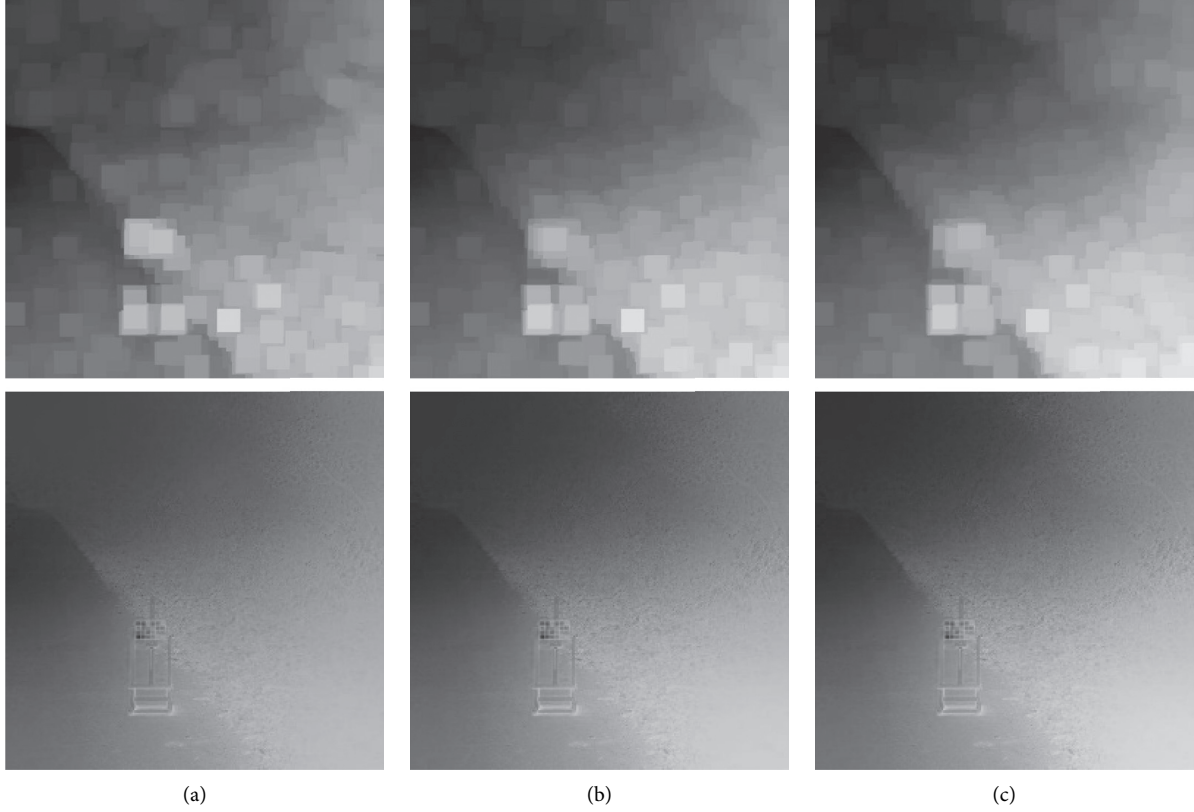


FIGURE 4: Comparison of the transmission before and after the guided filter. (a) Red channel. (b) Green channel. (c) Blue channel.

performance to the state-of-the-art algorithms including DCP [13], Meng et al.'s [16], Sulami et al.'s [17], Berman et al.'s [18], IDE [19], DehazeNet [26], and AOD [30] from both subjective and objective aspects. According to the experimental analysis, we set the window size as $s_l = 31$ in this study.

Image evaluation has extremely important application value in image dehazing, image compression, video compression, and other fields. Image evaluation methods are mainly divided into subjective evaluation and objective evaluation. Among them, subjective evaluation refers to the perceptual judgement based on observers' subjective feelings and is widely used in neural networks, image recognition, and other fields; however, it is susceptible to many capricious factors, such as personal preference, and is relatively ex parte.

To solve the problem of differentiation caused by subjective factors, it is necessary to formulate a standard for image quality evaluation (QA), that is, to use a specific mathematical model to quantify the difference between each image, to objectively evaluate the image quality.

We adopted the image visibility evaluation criterion and image structure similarity evaluation criterion using the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and natural image quality evaluator (NIQE) evaluation indicators to conduct a comprehensive evaluation and analysis of the aforementioned eight main dehazing algorithms. The PSNR uses the mean square error (MSE) to calculate the different values of the image pixels. The larger the value, the better the quality of the restored image. The calculation formula is shown in

$$\text{PSNR} = 10 \lg \left(\frac{(2^B - 1)^2}{\text{MSE}} \right). \quad (16)$$

Its MSE is calculated using

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2, \quad (17)$$

where I and K refer to a hazy image of size $m \times n$ and a processed image, respectively. The pixel value is represented by a B-bit binary number. In this study, the image pixel domain is $[0, 255]$; therefore, $B = 8$. The MSE represents the comparison result of the pixel-by-pixel difference between two images.

The structural similarity formula is implemented via brightness comparison formula (18), contrast comparison formula (19), and structure comparison formula (20):

$$l(i, j) = \frac{2\mu_i\mu_j + C_1}{\mu_i^2 + \mu_j^2 + C_1}, \quad (18)$$

$$c(i, j) = \frac{2\sigma_i\sigma_j + C_2}{\sigma_i^2 + \sigma_j^2 + C_2}, \quad (19)$$

$$s(i, j) = \frac{\sigma_{ij} + C_3}{\sigma_i\sigma_j + C_3}, \quad (20)$$

where μ_i and μ_j represent the mean values of i and j and σ_i^2 and σ_j^2 are the variances of i and j , respectively. σ_{ij} is the

covariance of i and j ; $C_1 = (k_1 L)^2$ and $C_2 = (k_2 L)^2$ are constants with default values $k_1 = 0.01$ and $k_2 = 0.03$ to avoid the case where the denominator is zero during the calculation. Equation (21) shows how the SSIM operates:

$$\text{SSIM}(i, j) = [l(i, j)^\alpha \cdot c(i, j)^\beta \cdot s(i, j)^\gamma]. \quad (21)$$

Considering $\alpha = \beta = \gamma = 1$, the calculation formula of SSIM can be obtained as

$$\text{SSIM}(i, j) = \frac{2\mu_i\mu_j + C_1}{\mu_i^2 + \mu_j^2 + C_1} \cdot \frac{2\sigma_{ij} + C_2}{\sigma_i^2 + \sigma_j^2 + C_2}. \quad (22)$$

SSIM has symmetry, which implies that $\text{SSIM}(i, j)$ is equivalent to $\text{SSIM}(j, i)$. The domain of SSIM is $[0, 1]$. The larger the value, the smaller the difference between the dehaze image and the original image. Images with large values show better integrity and a more reasonable structure. The evaluation criterion can also be used as an evaluation standard for the SR technology.

The NIQE algorithm evaluates the image quality by calculating the distance between the multivariate Gaussian (MGV) feature parameters extracted from the dehazing images and the pretrained model parameters, as shown in

$$D_{(v_1, v_2, \Sigma_1, \Sigma_2)} = \sqrt{(v_1 - v_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2)}, \quad (23)$$

where v_1 and v_2 are the mean vectors of the natural MGV model and Σ_1 and Σ_2 are the covariance matrices of the distorted image's MGV model.

The characteristic of NIQE is that no haze-free image is required as a reference. The smaller the value is, the smaller the gap between the haze-free image and the pretrained model is and the higher the image quality is. Therefore, this paper takes it as the evaluation index of real-world images.

4.1. Experiment on the O-Haze Dataset and NTIRE Dataset.

Substantial experiments were performed on the entire O-Haze and NTIRE datasets. The O-Haze dataset contains 45 paired hazy and haze-free images, whose hazy conditions are manufactured by a professional haze-making machine. The NTIRE dataset includes 35 pairs of outdoor nonhomogeneous hazy images. Part of the defogging results (with a size of 512×512) is shown in Figure 5 to facilitate comparison and analysis.

For the first image in the O-Haze dataset, the results of the DCP, Fattal, and AOD methods present an overall bluish hue; the Berman algorithm has a better foreground effect, but the distant details are faded, while our method, IDE, and the Meng algorithm retain clear distant details and show the best restoration effect; the IDE algorithm significantly increases the brightness. For the 11th image in the O-Haze dataset, all the algorithms achieved an insignificant dehazing effect on distant trees but our method outperformed the others on colour restoration in close-range processing. For the 12th image in the O-Haze dataset, our method achieved the most significant dehazing effect and the result is close to the real scene, whereas the results of the Fattal and AOD

methods are overall bluish and DehazeNet has a poor dehazing effect. For the 31st image in the O-Haze dataset, our algorithm has no obvious effect on the processing of the branches in the upper left corner, the overall colour of the Fattal algorithm is yellow, and the image of the Berman algorithm is distorted. For the 33rd image in the O-Haze dataset, all traditional algorithms have bright strips on the left side of the image and the effects of the two neural network algorithms are not obvious. Our algorithm and IDE algorithm achieved the best effect; for the 41st picture in the O-Haze dataset, our algorithm and Berman's algorithm are better than other algorithms in processing the sky area. For the 2nd image in the NTIRE dataset, our algorithm has the most thorough dehazing effect, the Berman algorithm has severe colour distortion, and the results of all other algorithms are bluish. For the 7th image in the NTIRE dataset, all algorithms have a poor dehazing effect owing to the heavy fog concentration, especially in the middle area above the image, but our algorithm and IDE algorithm have a better dehazing effect than other algorithms in the near field. For the 9th image in the NTIRE dataset, the DehazeNet algorithm does not have a significant dehazing effect. Our algorithm has the best defogging, but the hue is darker, and the other algorithms have different degrees of bluishness. In summary, our algorithm has certain advantages in terms of the subjective dehazing effect.

The results of the objective evaluation indexes PSNR and SSIM are listed in Table 1. As can be observed from the table, our algorithm has certain advantages over the other seven methods in both evaluation indexes. Most SSIM of our algorithm is optimal in most images in the O-Haze dataset and suboptimal in the NTIRE dataset, indicating that the gap between the defogging image and the original image is small and the structural integrity is strong. The DCP algorithm has a higher value when processing the 33rd image in the O-Haze dataset; however, from the perspective of the subjective defogging effect of the image, the brightness of the defogging image of DCP and the other four algorithms is low, whereas our algorithm significantly improves the overall brightness of the image, solves the problem of the bluish hue of dark channel dehazing, and is more in line with the characteristics of the real fog-free image. Although in the seventh image picture in the NTIRE dataset, the PSNR and SSIM of our algorithm are not the highest, the subjective feeling effect is better than that of other algorithms, which also shows that there is still room for improvement where the haze is particularly thick and efforts are required.

4.2. Experiment on Real-World Images.

To further verify the effectiveness of the algorithm in this study, real-scene images were selected for the experiment. The results are shown in Figure 6. The size of the foggy image was 512×512 . As shown in Figure 6, when processing the first and third images, the Meng, Fattal, and Berman algorithms exhibit serious colour distortion problems in the sky area, while the algorithm in this study has an ideal processing effect. For the second image, the detailed processing of the Fattal algorithm in the building part is not obvious, and the sky area shows a dark



FIGURE 5: Continued.

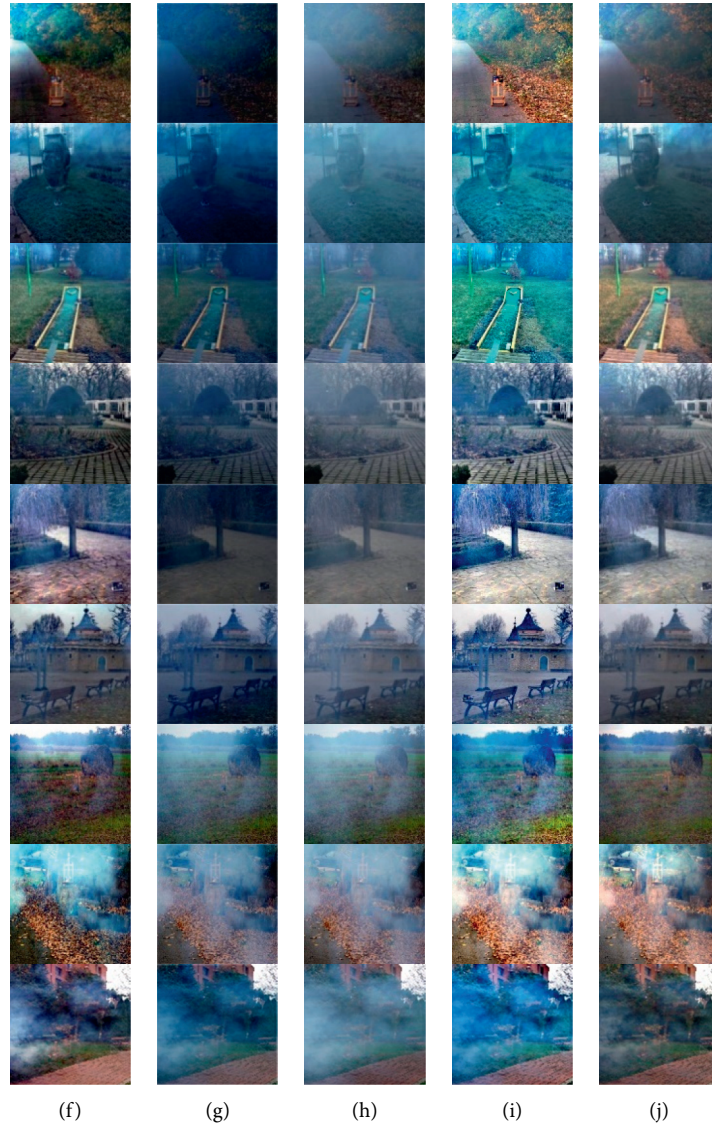


FIGURE 5: Output examples from the O-HAZE dataset and NTIRE dataset in the proposed method to compare with SOTA methods. (a) Original. (b) The fact. (c) DCP. (d) Meng. (e) Fattal. (f) Berman. (g) AOD. (h) DehazeNet. (i) IDE. (j) Ours.

hue. All the algorithms have no obvious effect on the sky region of the third image, and our algorithm and Fattal algorithm have an ideal effect in close view. For the fourth image, the colour of the Berman algorithm turns green, while AOD's processing and DehazeNet's processing of distant buildings lose details. The IDE algorithm can significantly improve brightness when processing real-world images. Compared with the DCP algorithm, the algorithm in this study is more natural in image defogging and has a higher colour contrast, which can realise image defogging technology in a variety of different scenes.

The NIQE values are shown in Table 2. It can be seen that in the images of road, urban area, and urban rail, IDE algorithm obtains low parameter values due to overexposure of the images, but obviously, it cannot be applied to all the images. As for image parameters of urban area images, IDE obtains the highest parameter values among the 8

algorithms. The proposed algorithm is more consistent with people's subjective feelings in the road image, and the detail processing of the distant houses in the city rail image are better than the IDE algorithm. The proposed algorithm obtains relatively low values in the processing of urban images and is second only to the DehazeNet algorithm in the housing images. The reason is that there exists colour distortion in the upper left corner of the image. But, the DehazeNet algorithm has a poor haze-removing effect in terms of subjective perception. Comprehensively, we believe that our algorithm preserves more original colour information in the image dehazing process and the effect is better.

From the above experimental comparison, it can be observed that the proposed algorithm can achieve a finer effect in the processing of image details and the resulting images are more realistic, providing people with a better visual experience. The parameters PSNR and SSIM are

TABLE 1: SSIM/PSNR performance of different methods on the O-HAZE dataset and NTIRE dataset.

Index	Metrics	DCP	Meng	Fattal	Berman	AOD	DehazeNet	IDE	Our
O-Haze -01	PSNR	17.0322	20.0174	17.0816	17.5378	11.6131	13.7935	18.3342	23.4737
	SSIM	0.5734	0.6847	0.6365	0.6477	0.4740	0.3858	0.6763	0.7189
O-Haze -11	PSNR	19.0962	20.5596	18.1265	17.1912	11.5011	13.3388	19.8388	22.1190
	SSIM	0.5108	0.6214	0.5275	0.5733	0.3655	0.6232	0.7141	0.6111
O-Haze -12	PSNR	20.5862	22.5818	18.4336	20.6551	14.5371	14.3256	18.2861	23.8214
	SSIM	0.7250	0.7441	0.6732	0.7396	0.6286	0.1660	0.7336	0.7423
O-Haze -31	PSNR	20.9985	20.2618	21.1603	21.1059	15.6592	19.0675	22.2526	28.1259
	SSIM	0.7381	0.7429	0.7426	0.7381	0.7005	0.6599	0.7529	0.8289
O-Haze -33	PSNR	25.0917	22.6296	18.3985	19.2103	15.8634	17.9752	16.1411	26.0924
	SSIM	0.7545	0.6529	0.6885	0.5917	0.6765	0.5841	0.6058	0.7213
O-Haze -41	PSNR	19.1044	19.2961	17.1685	21.1482	14.8864	21.394	22.0644	22.4768
	SSIM	0.7054	0.7196	0.7490	0.7883	0.7257	0.5718	0.8063	0.7990
NTIRE -02	PSNR	17.1174	17.9890	17.0191	14.7812	9.5366	10.1905	19.2308	19.6378
	SSIM	0.6846	0.7809	0.8253	0.5975	0.7959	0.7722	0.8509	0.7995
NTIRE -07	PSNR	20.8645	20.3936	18.8785	15.7877	11.1128	10.9838	14.9667	20.6927
	SSIM	0.7784	0.8006	0.7396	0.6092	0.7126	0.6590	0.7141	0.7982
NTIRE -09	PSNR	16.4644	16.0113	15.6139	17.0456	9.0514	7.1131	16.6816	20.3326
	SSIM	0.5138	0.5621	0.5969	0.6941	0.6197	0.5017	0.6608	0.6637

Our proposed method outperforms the others. The bold values represent the best performance of all candidate algorithms.

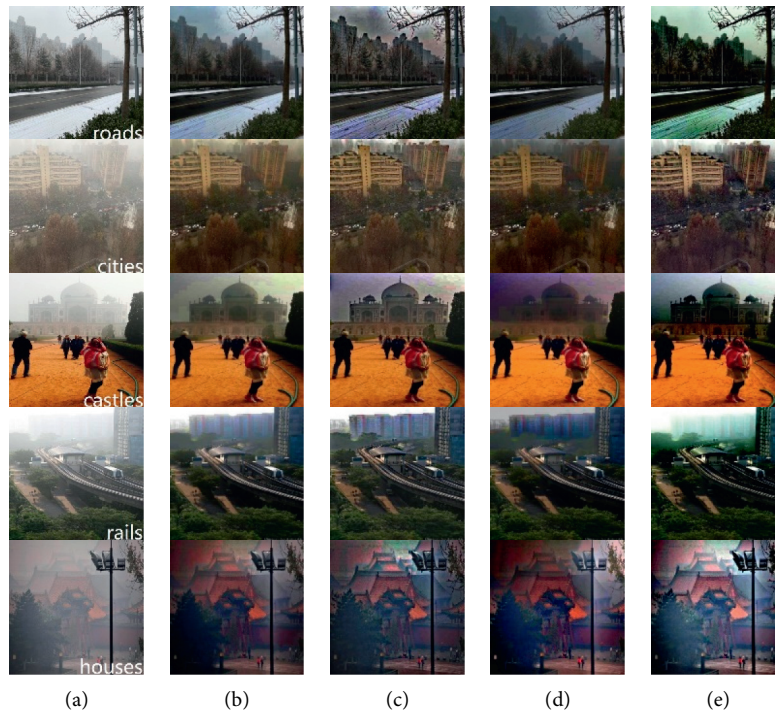


FIGURE 6: Continued.



FIGURE 6: Output examples from real-world images in the proposed method to compare with SOTA methods. (a) Original. (b) DCP. (c) Meng. (d) Fattal. (e) Berman. (f) AOD. (g) DehazeNet. (h) IDE. (i) Ours.

TABLE 2: NIQE evaluation of real-world data sets.

	DCP	Meng	Fattal	Berman	AOD	DehazeNet	IDE	Ours
Roads	2.7486	3.0143	2.7038	2.8870	2.8608	2.7627	2.4398	2.8246
Cities	2.5885	2.6136	2.6534	2.6203	2.9298	2.8441	3.1798	2.4755
Castles	2.5696	2.6897	2.6576	2.6609	2.5347	2.6390	2.5017	2.5908
Rails	3.7564	3.8570	3.6773	3.6989	3.8285	3.6112	2.5602	3.7711
Houses	2.3889	2.3241	2.4281	2.7754	2.7132	2.2252	3.3244	2.3168

The bold values represent the best performance of all candidate algorithms.

higher than those of the other algorithms in most cases. However, the algorithm proposed in this study has a poor effect in thick haze areas. This is because, under the influence of large white areas, the significance of the regional similarity value is lost, which leads to excessive brightness exposure in such regions after dehazing, resulting in regional white blocks.

5. Conclusion

In this paper, a dehazing algorithm for heterogeneous images based on the optimisation of regional atmospheric ambient light was proposed to solve the problem of image colour distortion and the overall bluish and somber hue, which exists in the dark channel prior algorithm. In this method, the regional mean value of the RGB three-layer dark channel in a foggy image was solved and the corresponding regional atmospheric ambient light value was obtained by linear stretch correction. Subsequently, according to the transmittance obtained by the dark channel prior theory, the V channel of the foggy image is considered as the guiding graph for guiding filtering according to different channels to approach the real image transmittance and the atmospheric

scattering model is used for defogging processing. The experimental results show that the image dehazing optimised by the proposed algorithm is superior to most classical dehazing algorithms in both subjective vision and objective quantification indexes and the problem of the dark hue is well solved. However, the algorithm in this study should traverse the entire image when calculating the atmospheric ambient light value, leading to a long execution time of the algorithm. To achieve better defogging efficiency, this study will continue to explore a better solution.

Data Availability

The research data used to support the findings of this study are available from the corresponding author only under license and so cannot be made freely available. Requests for access to these data should be made to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was mainly supported by the Beijing Municipal Education Commission Research Foundation under Grant KM202111417008, the Academic Research Projects of Beijing Union University under Nos. ZK90202104 and ZB10202004, and the Beijing Municipal Natural Science Foundation under Grant L191006 and partly supported by Beijing Union University Students' Science and Technology Innovation Project (no. 20213065).

References

- [1] D. L. Wang and T. Y. Zhang, "Review and analysis of image defogging algorithm," *Journal of Graphics*, vol. 41, no. 6, pp. 861–870, 2020.
- [2] H. W. Jiang, Z. Yang, X. Zhang, and Q. L. Dong, "Research progress of image dehazing algorithms," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 51, no. 4, pp. 1169–1181, 2021.
- [3] C. X. He, "Research progress on image dehazing algorithm," *Modern Computer*, vol. 700, no. 28, pp. 48–52, 2020.
- [4] E. H. Land and J. J. McCann, "Lightness and Retinex theory," *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1–11, 1971.
- [5] Z. L. Ma and J. Wen, "Single-scale Retinex sea fog removal algorithm fused the edge information," *Journal of Computer-Aided Design & Computer Graphics*, vol. 27, no. 2, pp. 217–225, 2015.
- [6] M. Zheng, G. Qi, Z. Zhu, Y. Li, H. Wei, and Y. Liu, "Image dehazing by an artificial image fusion method based on adaptive structure decomposition," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 8062–8072, 2020.
- [7] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, and N. Mazur, "A novel fast single image dehazing algorithm based on artificial multiexposure image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–23, 2021.
- [8] Q. L. Shu, C. S. Wu, Q. X. Zhong, and R. W. Liu, "Alternating minimization Algorithm for Hybrid Regularized Variational Image Dehazing," *Optik - International Journal for Light Electron Optics*, vol. 185, 2019.
- [9] W. Wang, S. X. Hu, and Z. Pei, "Underwater image enhancement algorithm based on dense fusion GAN," *Journal of Xi'an Polytechnic University*, pp. 1–8, 2021.
- [10] S. K. Nayar and S. G. Narasimhan, "Vision in bad weather," *IEEE Int'l Conf. Computer Vision*, vol. 38, no. 5, pp. 1181–1186, 1999.
- [11] X. L. Jin, W. Zhang, and L. F. Liu, "Image defogging algorithm based on guided filtering and adaptive tolerance," *Journal on Communications*, vol. 41, no. 5, pp. 31–40, 2020.
- [12] Y. N. Wang and X. B. Yang, "An improved algorithm of image dehazing based on sky region recognition," *Journal of Detection and Control*, vol. 042, no. 2, pp. 71–78, 2020.
- [13] K. M. Kaiming He, J. Jian Sun, and X. O. Xiaou Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.
- [14] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.
- [15] Z. Zhu, Y. Luo, H. Wei et al., "Atmospheric light estimation based remote sensing image dehazing," *Remote Sensing*, vol. 13, no. 13, p. 2432, 2021.
- [16] G. Meng, Y. Wang, J. Y. Duan, S. M. Xiang, and C. H. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *Proceedings of the 2013 IEEE International Conference on Computer Vision*, Sydney, Australia, December 2013.
- [17] M. Sulami, I. Glatzer, R. Fattal, and M. Werman, "Automatic recovery of the atmospheric light in hazy images," in *Proceedings of the IEEE International Conference on Computational Photography*, Santa Clara, CA, USA, May 2014.
- [18] D. Berman, T. Tali, and A. Shai, "Non-local image dehazing," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, July 2016.
- [19] M. Ju, C. Ding, W. Ren, Y. Yang, D. Zhang, and Y. J. Guo, "IDE: image dehazing and exposure using an enhanced atmospheric scattering model," *IEEE Transactions on Image Processing*, vol. 30, pp. 2180–2192, 2021.
- [20] M. Y. Ju, C. Ding, W. Q. Ren, and Y. Yang, "IDBP: Image dehazing using blended priors including non-local, local, and global priors," *IEEE Transactions on Circuits and Systems for Video Technology*, 2021.
- [21] W. Wang, X. Yuan, X. Wu, and Y. Liu, "Fast image dehazing method based on linear transformation," *IEEE Transactions on Multimedia*, vol. 19, no. 6, pp. 1142–1155, 2017.
- [22] M. Y. Ju, C. Ding, Y. J. Guo, and D. Y. Zhang, "IDGCP: image dehazing based on gamma correction prior," *IEEE Transactions on Image Processing*, vol. 29, 2019.
- [23] M. Y. Ju, C. Ding, and Y. J. Guo, "VROHI: Visibility recovery for outdoor hazy image in scattering media," *IEEE Photonics Journal*, vol. 12, 2020.
- [24] M. Y. Ju, C. Ding, D. Y. Zhang, and Y. J. Guo, "BDPK: Bayesian dehazing using prior knowledge," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, pp. 2349–2362, 2018.
- [25] Y. Chen, H. G. Guo, and Y. P. Ai, "Single image dehazing method based on multi-scale convolution neural network," *Acta Optica Sinica*, vol. 39, no. 10, 2019.
- [26] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: an end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.
- [27] J. J. Li, G. H. Li, and H. Fan, "Image dehazing using residual-based deep CNN," *IEEE Access*, vol. 6, p. 1, 2018.
- [28] X. Qin, Z. L. Wang, Y. C. Bai, X. D. Xie, and H. Z. Jia, "FFA-net: feature fusion attention network for single image dehazing," in *Proceedings of the Association for the Advance of Artificial Intelligence*, vol. 34, no. 7, pp. 11908–11915, 2019.
- [29] R. W. Liu, W. Q. Yuan, X. Q. Chen, and Y. X. Lu, "An enhanced CNN-enabled learning method for promoting ship detection in maritime surveillance system," *Ocean Engineering*, vol. 235, Article ID 109435, 2021.
- [30] B. Y. Li, X. L. Peng, Z. Y. Wang, J. Z. Xu, and D. Feng, *An All-In-One Network for Dehazing and beyond*, Institute of Electrical and Electronics Engineers, Manhattan, NY, USA, 2017.
- [31] H. Dong, J. Pan, L. Xiang et al., *Multi-Scale Boosted Dehazing Network with Dense Feature Fusion*, Institute of Electrical and Electronics Engineers, Manhattan, NY, USA, 2020.
- [32] Y. J. Shao, L. Li, W. Q. Ren, C. X. Gao, and N. Sang, *Domain Adaptation for Image Dehazing*, Institute of Electrical and Electronics Engineers, Manhattan, NY, USA, 2020.
- [33] Y. W. Pang, J. Nie, J. Xie, J. G. Han, and X. L. Li, "BidNet: binocular image dehazing without explicit disparity estimation," in *Proceedings of the 2020 IEEE/CVF Conference on*

Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, June 2020.

- [34] H. Y. Wu, Y. Qu, S. Lin et al., “Contrastive learning for compact single image dehazing,” in *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, Tennessee, June 2021.

Research Article

Evaluation of Blended Oral English Teaching Based on the Mixed Model of SPOC and Deep Learning

Yanli Hui 

Faculty of Foreign Languages and Business, Jiaozuo Normal College, Jiaozuo 454001, China

Correspondence should be addressed to Yanli Hui; 1295004004@jzsz.edu.cn

Received 8 September 2021; Revised 14 October 2021; Accepted 26 October 2021; Published 8 November 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Yanli Hui. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the deep integration of “internet + education” and the continuous advancement of education reform, blended teaching has become the main method of university education reform. Blended education combines the advantages of traditional education and online education to complement each other. It not only takes advantage of the flexibility and autonomy of online education but also retains the benefits of emotional communication between teachers and students in offline education. With the increase in practical exploration of blended teaching in universities, teaching evaluation is an important part of teaching, and blended teaching evaluation should also attract attention. The purpose of this paper is to study the mixed oral English teaching evaluation based on the mixed mode of SPOC and deep learning. On the basis of analyzing the teaching design principles of the mixed mode of SPOC and deep learning and the principles of constructing the teaching evaluation after half a semester of teaching investigations conducted by the two classes of English majors, the impact of the SPOC and deep learning mixed teaching mode on students’ spoken English was studied through the method of covariance analysis. The experimental results show that the mixed teaching mode of SPOC and deep learning has been able to fully stimulate students’ interest in oral English learning and improve students’ oral English ability, critical thinking of students, ability to solve problems, group cooperation, and effective communication. Self-directed learning and self-reflection have all had a positive impact.

1. Introduction

With the acceleration of the globalization of the world economy, English has become a commonly used language throughout the world and has played a great role in promoting economic, cultural, academic, and political exchanges and cultural exchanges between various ethnic groups [1, 2]. Spoken language is its external form of expression, which has attracted great attention from all countries in the world. However, for a long time, in our country, oral English teaching often adopts the teaching method of “students learn passively under the guidance of teachers, and the interaction between teachers and students is weakened.” Due to the limited resources of textbooks and the lack of a real language environment, students’ oral English expression skills are generally weak, and phenomena such as “high scores and low energy” and “dumb English” still exist. In order to change this situation, it is urgent to

make full use of modern information technology and adjust the college English teaching mode [3, 4]. English is a West Germanic language branch. It was first used by medieval Britain and has become the most widely used language in the world because of its vast colonies. The Anglo tribe, the ancestor of the British, is one of the Germanic tribes that later migrated to the island of Great Britain, known as England. Both names come from Anglia on the Baltic peninsula. The language is closely related to Frisian and lower Saxon. Its vocabulary is influenced by other Germanic languages, especially Nordic, and is largely written in Latin and French.

In recent years, many scholars have conducted research on mixed nurse teaching and have achieved good results. Some scholars believe that it is the use of the Pan-Asian SPOC concept and the in-depth integration of traditional classroom teaching theories at home and abroad. Besides, SPOC perfectly adapts to the exclusiveness of elite

universities and the values of pursuing high achievements; SPOC model has low cost and can be used to generate income, which provides a sustainable development model of MOOC; SPOC redefines the role of teachers and innovates the teaching model; SPOC puts more emphasis on students' complete and in-depth learning experience, which is conducive to improving the completion rate of the course. Based on a systematic comparison of the differences between MOOC and SPOC, they elaborated on several typical practical applications of SPOC at home and abroad. The case explored the design of a hybrid classroom teaching model based on the Pan-Asian SPOC. The teaching theory was fully used in the platform, and the model including front-end analysis, curriculum design, curriculum organization, and teaching "four-core" evaluation was analyzed in detail [5]. In addition, some scholars have organically integrated classroom teaching and online learning by constructing a SPOC hybrid teaching model and developed a flipped classroom. From "teaching as the center" to "learning" "centered," the new classroom assessment mode adopts a variety of assessment methods to effectively ensure the quality of classroom teaching according to the learning characteristics, professional background, and learning requirements of students, so as to promote students' independent learning [6]. The results provide theoretical guidance for the research of this article.

Based on the literature review, this article will explore mixed oral English based on the SPOC and deep learning mixed mode by comparing and analyzing the oral English situation of the students in the experimental class and the control class before and after the experiment, as well as the evaluation of the mixed mode teaching. Deep learning technology can be used to verify the effectiveness of the teaching. Deep learning has made many achievements in search technology, data mining, machine learning, machine translation, natural language processing, multimedia learning, voice, recommendation and personalization technology, and other related fields. Deep learning makes machines imitate human activities such as audiovisual and thinking, solves many complex pattern recognition problems, and makes great progress in artificial intelligence-related technologies.

2. Research on the Evaluation of Blended Oral English Teaching Based on the Hybrid Model of SPOC and Deep Learning

2.1. The Teaching Design Principles of the Hybrid Model of SPOC and Deep Learning

2.1.1. *Teacher-Led and Student-Oriented Principles.* The traditional classroom teaching model in our country has long tended to be teacher-centered, focusing on giving play to the leadership role of teachers in the classroom. However, under such an education system, the initiative of students is constantly weakening, which is not conducive to the enthusiasm and creativity of students. The blended learning model not only reflects the leading role of teachers in education but also has a complete impact on the subjectivity of students. The design of the blended learning model should

adhere to the student-centered principle. In the education process, students are the main body of learning, and teachers are only assistants. Teachers need to design logical courses, encourage students to learn actively and continuously, and cultivate their independent learning ability and innovative consciousness. Through the blended learning model, the teacher-led and student-based subjects are deeply integrated, and finally, an excellent educational effect is achieved [7, 8]. One of the important characteristics of teaching different from other cognition or learning activities is that students' cognition is carried out under the guidance of teachers. Teachers have received special education and training. They understand the purpose of teaching and master the contents and methods of teaching. Therefore, it is objectively inevitable for teachers to play a leading role. In the process of teaching, students are the main body of cognitive or learning activities. Therefore, students' initiative, enthusiasm, and creativity are also the necessary conditions for successful teaching.

2.1.2. *The Principle of Integrity.* Blended learning is a teaching method that combines the advantages of traditional classroom teaching and online learning. This way of design needs to include front-end analysis, learning process design, learning evaluation design, and many other links. Each link is interrelated and cannot be separated. Only when all links work together can the overall function be completed [9, 10].

2.1.3. *Evaluative Principle.* As the name suggests, online education is a teaching method based on the network. Through the network, students and teachers can carry out teaching activities even if they are thousands of miles apart; in addition, with the help of network courseware, students can study anytime and anywhere, which really breaks the restrictions of time and space. For workplace people with busy work and uncertain learning time, network distance education is the most convenient way of learning.

All education models are inseparable from an objective evaluation mechanism. Evaluation mechanisms play an important role in the design of complex learning models. The traditional education model often regards the student's academic performance as the only criterion for evaluating students, while the mixed learning model requires the use of multiple evaluation methods. When evaluating students, not only the academic performance of the students must be considered but also the sexual evaluation. Formative evaluation helps students maintain self-confidence, maintain enthusiasm for learning, and promote student learning. Therefore, when designing a hybrid learning model, the principle of evaluation should always be considered, and the model should be continuously improved based on the results of evaluation and feedback [11, 12].

2.2. SPOC and Deep Learning Hybrid Model Teaching Evaluation Construction Principles

2.2.1. *Principle of Consistency with the Target.* Indicators are expressed as specific, behavioral, and functional goals. These

goals must be fully reflected and highly consistent with training goals or daily management standards. Consistency with the target refers to the consistency of the indicators in the system. The SPOC scoring index system based on the perspective of deep learning avoids the combination of two conflicting indicators, so as not to confuse the evaluator and fail to make an accurate evaluation of the SPOC. The principle of unity of objectives means that the more the contribution of all departments and members in the organization is conducive to the realization of the organization's objectives, the more reasonable the organization's structure is. Organization is a cooperative system created to achieve goals. The establishment of an organization is the means to achieve goals. The common goal is the objective basis for the establishment and existence of an organization. Without a common goal, it is impossible to establish an organization. Even if it is established temporarily, it is impossible to survive for a long time. Only with clear and consistent goals can all departments and members of the organization have the basis for cooperation and common direction of action; otherwise, there can be neither intention nor action of cooperation. The common goal is also the objective basis for improving and developing the organization. The improvement and development of the organization must be based on the realization of common goals; otherwise, it will destroy the organization.

2.2.2. Concise and Scientific Principles. The evaluation of SPOCs should follow scientific principles and objectively and honestly reflect the characteristics and current situation of SPOCs from a detailed learning perspective. This enables the evaluation to truly reflect the current problems of the university's SPOC curriculum, the indicators are not easy to be too complicated, convenient for students to observe and measure, combined with the characteristics of deep learning, improve the curriculum, promote the development of university SPOCs, and enable deep learning in the education team, and the emergence of individual students cultivates practical and creative talents who can solve complex problems.

2.2.3. Principles from the Perspective of Deep Learning. When determining the dimensions of the scoring indicators and the specific elements of the scoring indicators, we should start from the perspective of deep learning, comprehensively consider the factors that affect the deep learning of SPOCs, analyze the data, and extract the grading index from the perspective of deep learning to examine its accurate expression so that evaluators have accurate positioning and understanding of indicators when evaluating courses with reference to standards. The calculation involved in generating an output from an input can be represented by a flowchart: the flowchart is a graph that can represent the calculation. In this graph, each node represents a basic calculation and a calculated value, and the calculated results are applied to the values of the child nodes of this node. Consider such a calculation set, which can be allowed in each node and possible graph structure, and define a function

family. The input node has no parent nodes, and the output node has no child nodes.

3. Experiment

3.1. Subjects. In order to verify the feasibility and effectiveness of the mixed teaching model designed in this article, this article has conducted a half-semester teaching survey on two classes of non-English majors in a university and studied the mixed teaching of SPOC and deep learning through the method of covariance analysis. We study the impact of this model on student performance and the development of oral English and analyze the effectiveness of this model from the perspective of teachers and students. Before the implementation of the teaching model, this article first investigated the prespeaking test scores of non-English majors as a comparison for follow-up research. In the following practice, this article will use one class as the control class and the other as the experimental class. The two classes have the same learning foundation, and there is no difference in the ratio of men to women, the content of the teaching materials, the teaching progress, and the oral level.

Covariance analysis is also called "covariate (number) analysis." Analysis of covariance is essentially the extension and expansion of analysis of variance. The basic principle is to combine linear regression with analysis of variance, adjust the average of each group and the experimental error term of the F -test, and test whether there is a significant difference between two or more adjusted averages, so as to control the influence of covariates (variables with a close regression relationship with dependent variables) that affect the experimental effect (dependent variables) in the experiment and are not controlled by legal person in the analysis of variance. Covariance is an overall parameter used to measure the "collaborative variation" between two variables, that is, the parameter of the interaction between two variables. The greater the absolute value of covariance, the greater the interaction between two variables.

3.2. Experimental Process

3.2.1. Before the Experiment. Before the experiment, the students in the control class and the experimental class received a three-week oral training on a regular basis, and students are organized to take pretraining exams. The preexamination paper has a total of 100 units, and each student takes about 6 minutes to take the exam. The test paper is divided into self-introduction parts. The second part is the keynote speech. The third part is the question and answer session.

3.2.2. In the Experiment. In the course of the experiment, the control class adopts the method of oral teaching with the help of multimedia courses. Teachers integrate resources for oral teaching planning during SPOC classroom preparation, and teachers guide students to use the internet and other channels to find resources for self-preparation. In the classroom, through collaborative learning and group

discussions, special lectures, and status presentations, the ability to use spoken English is cultivated. Teachers encourage students to participate in interactive discussions to improve their ability to use spoken English.

3.2.3. After the Experiment. After the experiment, the students in the control class and the experimental class will take the posttest exam. The postexperiment test consists of three parts. The first part is self-introduction, the second part is a keynote speech, and the third part is a question and answer session. The test time per life is 6 minutes.

3.3. Data Processing. In order to improve the reliability and validity of the experimental data, this article discusses the use of the covariance analysis method in SPSS 17.0 to statistically analyze the data scored by the two classes before and after the exam. The predicted results are analyzed as the covariance. Note that SPSS is the earliest statistical software in the world that uses a graphical menu-driven interface. Its most prominent feature is that the operation interface is very friendly, and the output results are beautiful. It shows almost all functions in a unified and standardized interface and uses the Windows window to show the functions of various data management and analysis methods, and the dialog box shows various function options. As long as users master certain Windows operation skills and master the principle of statistical analysis, they can use software to serve specific scientific research work. The regression model is used to verify the results of the pretest, and the oral teaching method based on the flipped classroom has a predictive effect on the teaching effect after the experiment; thus, the impact of the flipped classroom oral teaching method on its performance in the postexperiment evaluation is established. The calculation process of the covariance analysis method is as follows:

$$Y - X_2\beta = X_1\alpha + e. \quad (1)$$

According to formula (1), find the least square estimate of α :

$$\alpha = (X_1^T X_1)^{-1} X_1^T (Y - X_2\beta). \quad (2)$$

From $Y - X_2\beta = X_1\alpha + e$, find the least square estimate of β :

$$\beta = (X_2^T X_2)^{-1} X_2^T (Y - X_1\alpha). \quad (3)$$

Among them, $X_1\alpha$ is the variance analysis part of the model, and $X_2\beta$ is the regression part of the model.

4. Discussion

4.1. Comparison of Learning Interests. In order to verify the influence of this teaching model on students' learning interest, after the experiment, this article surveyed the students' oral learning interest in the two classes in the pretest questionnaire and compared the students' oral learning interest before and after the experiment.

According to Table 1 and Figure 1, it can be seen that the learning interest of the students in the control class has improved before and after the experiment, but the data changes are not significant, indicating that, after a semester of study, the students' interest in oral learning in the control class has not changed much. 36.777% of the students in the experimental class strongly agreed that the SPOC teaching model stimulated their interest in learning, and 46.77% of the students agreed that this teaching model enhanced their interest in learning, which was an increase of 16.67% from before the experiment. It can be seen that SPOC blended teaching has a very obvious effect on stimulating students' interest in learning, and the SPOC blended teaching model has a greater role in promoting students' interest in learning.

4.2. Comparison of Learning Effects. After a semester of teaching practice, this article compares and analyzes the oral learning effects of the experimental class and the control class based on the posttest questionnaire survey results and the oral test results. The situation is as follows.

According to Table 2 and Figure 2, it can be concluded that 14.28% of the control class think that their oral expression ability has been greatly improved, and 21% of the experimental class and 32.13% of the control class think that their oral proficiency has been improved. There are 47.67%, 17.88%, and 6.68% of the students in the control class and the experimental class who think that their oral expression skills have not been improved, and the proportion of the number of students in the experimental class is significantly larger than that in the control class.

According to Figure 3, the average score of the oral test of the experimental class is 81.76, and the average score of the control class is 70.35. Compared with the scores of the previous test, the scores of both classes have improved, and the scores of the experimental class have increased by approximately 12 points, the results of the control class increased by about 5 points, and the results of the experimental class improved even more, indicating that the mixed oral English teaching based on the mixed mode of SPOC and deep learning is effective.

4.3. Comparison of Deep Learning Capabilities. The change of deep learning ability is used as the main criterion to evaluate the mixed oral English teaching based on SPOC and deep learning. The method of evaluation is through the deep learning ability questionnaire before and after the teaching of the experimental group students and collecting data, comparing SPOC and deep learning. The relationship between the mixed mode and the difference of the students' deep learning ability before and after the teaching mode is not implemented. Note that cognition includes feeling, perception, memory, thinking, imagination, and language. Specifically, the process of acquiring or applying knowledge begins with feeling and perception. Feeling is the understanding of the individual attributes and characteristics of things. Traditionally, it refers to the cognitive process, which refers to the psychological process in which the human brain reflects the characteristics and relations of objective things in

TABLE 1: Changes in learning interest of the two classes before and after the experiment.

	Control class before the experiment (%)	Control class after the experiment (%)	Experimental class before the experiment (%)	Experimental class after the experiment (%)
Incompatible	3.58	3.67	3.23	0
Not very consistent	14.31	7.2	9.8	3.23
Basically accord with	46.42	46.33	46.77	16.57
Conform to	24.99	28.9	30.1	46.77
Very much in line	10.7	13.9	10.1	36.77

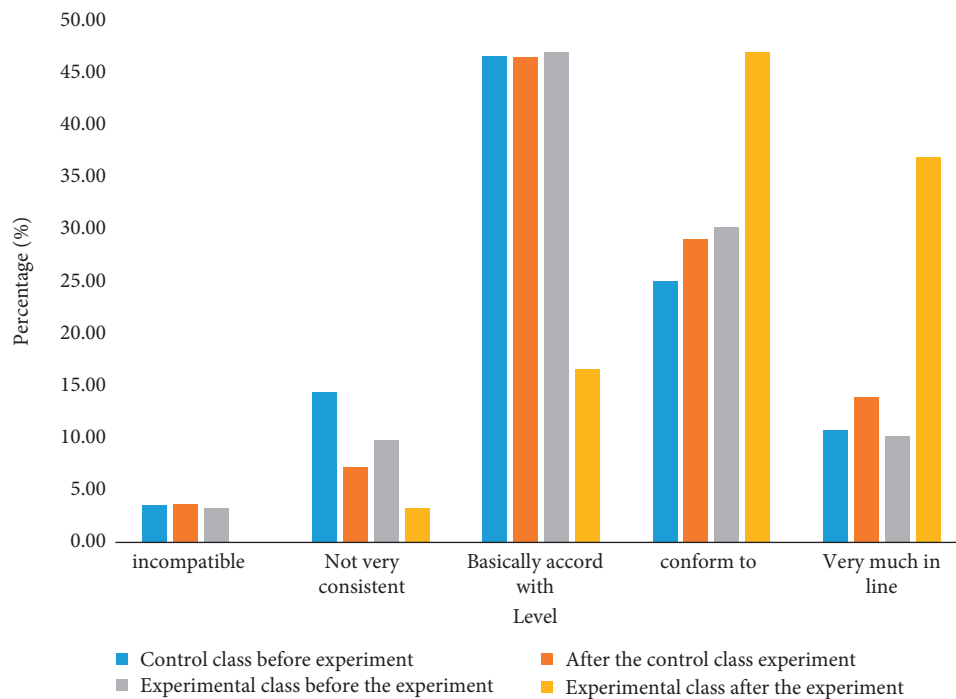


FIGURE 1: Changes in learning interest of the two classes before and after the experiment.

TABLE 2: Comparison of students' oral English improvement in two classes.

	Control class (%)	Experimental class (%)
Very much in line	14.28	21
Conform to	32.13	47.67
Basically accord with	28.56	21.33
Not very consistent	17.88	6.68
Incompatible	7.15	3.32

the form of perception, memory, and thinking. In cognitive psychology, information processing refers to the process in which individuals receive, encode, store, extract, and use information.

It can be seen from Table 3 and Figure 4 that the pretest average in the cognitive domain is 3.64, and the posttest average is 3.70; the pretest average in the interpersonal domain is 3.68, and the posttest average is 3.74; the pretest average in the personal domain is 3.81, and the posttest average is 3.84.

As can be seen from Table 4 and Figure 5, the cognitive domain correlation coefficient is 0.94, the t -value of the paired sample test is -5.264 , and the significance $p = 0.000 < 0.01$, showing a significant difference at the level of 0.01; the interpersonal domain correlation coefficient is 0.961, the t -value of the paired sample test is -5.352 , and the significance $p = 0.000 < 0.01$, showing a significant difference of the 0.01 level; the personal field correlation coefficient is 0.968, the t -value of the paired sample test is -4.310 , and the significance

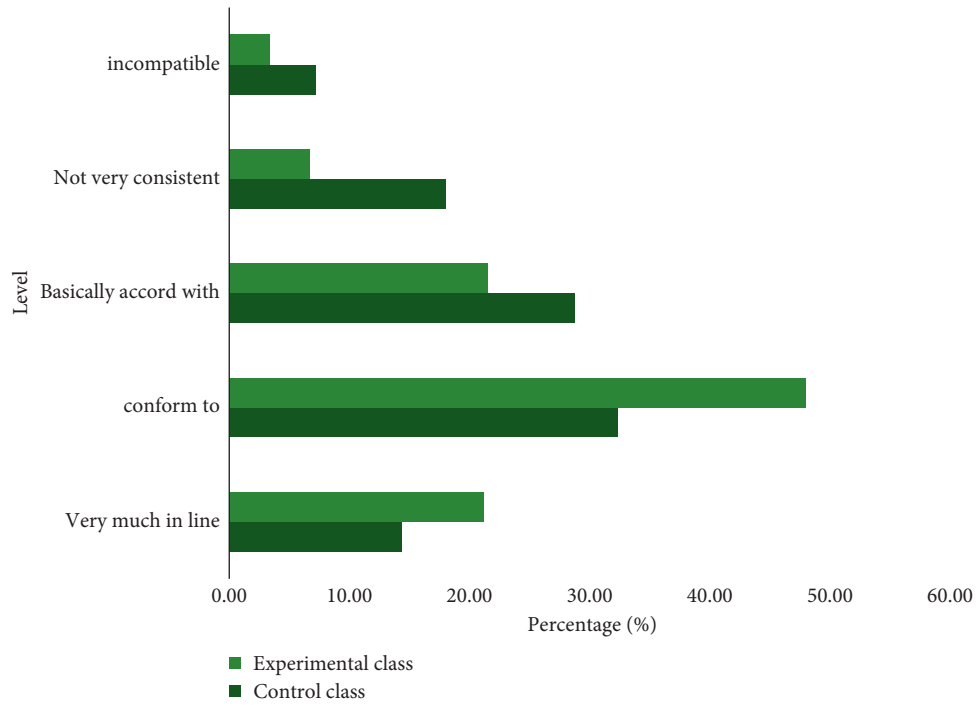


FIGURE 2: Comparison of students' oral English improvement in two classes.

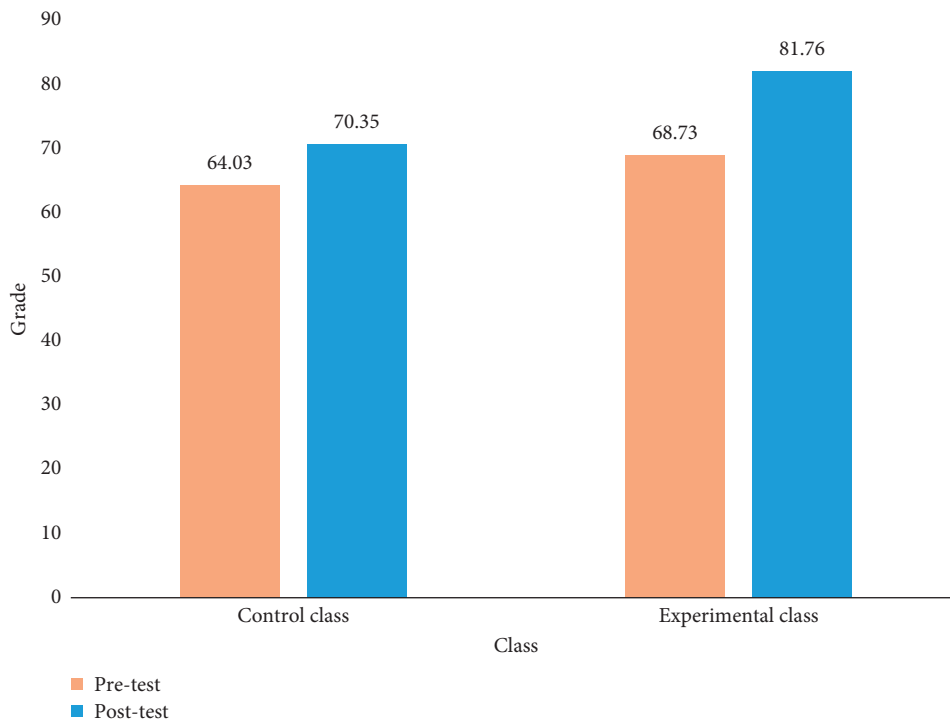


FIGURE 3: Comparison of two and a half oral test scores.

TABLE 3: Paired sample statistics.

		Average value	Standard deviation	Standard deviation
Pairing 1	Cognitive front	3.64	0.742	0.036
	Cognitive posttest	3.70	0.633	0.031
Pairing 2	Interpersonal pretest	3.68	0.968	0.045
	Interpersonal posttest	3.74	0.890	0.041
Pairing 3	Individual front	3.81	0.834	0.035
	Individual posttest	3.84	0.780	0.033

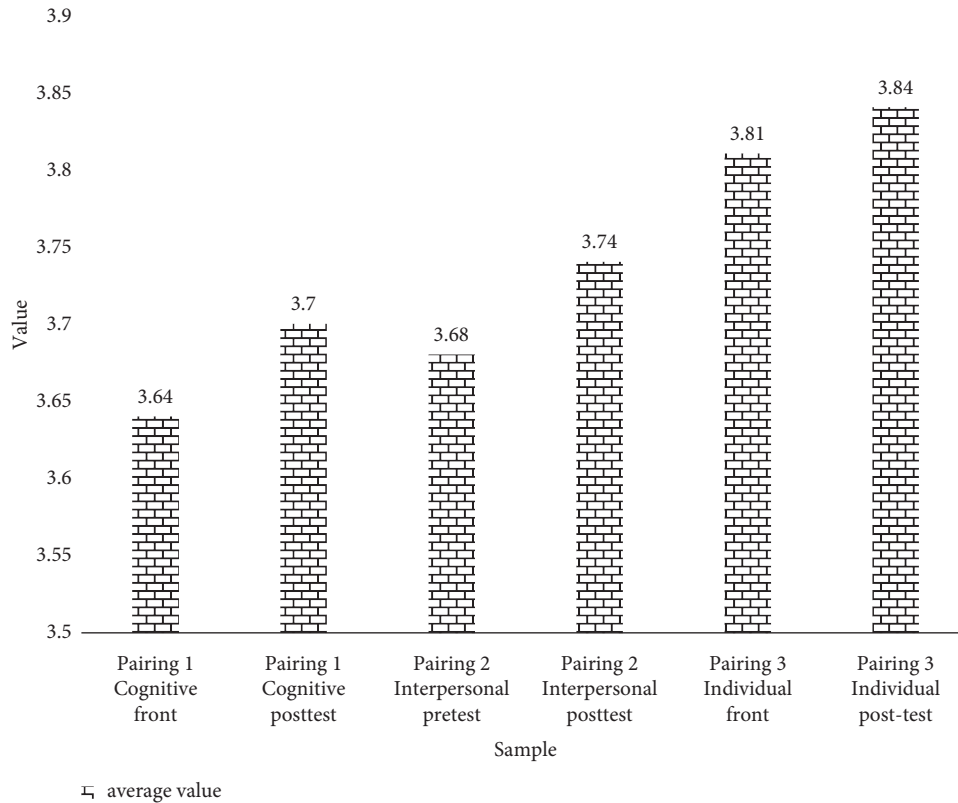


FIGURE 4: Paired sample mean.

TABLE 4: Paired sample testing and correlation.

		Average value	<i>t</i>	Correlation	Significance
Pairing 1	Cognitive front-cognitive posttest	-0.063	-5.264	0.940	0.000
Pairing 2	Interpersonal pretest-interpersonal posttest	-0.059	-5.352	0.961	0.000
Pairing 3	Individual front-individual posttest	-0.032	-4.310	0.968	0.000

$p = 0.000 < 0.01$, showing a significant difference at the level of 0.01; these data show that, after half a semester of teaching practice, the mixed oral English teaching based on the SPOC and deep learning mixed mode has a critical

thinking and responsible problem-solving ability for students, teamwork, and effective communication; independent learning and self-reflection have all had a positive impact.

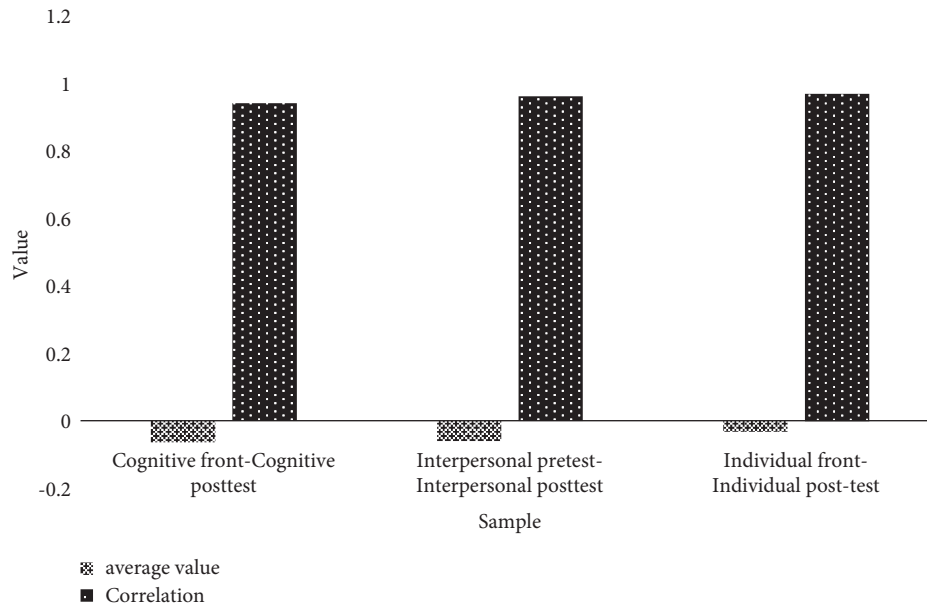


FIGURE 5: Paired sample mean and correlation.

5. Conclusions

The rapid development of education informatization has brought new opportunities and challenges for teachers. Network learning, mobile learning, and other means make students' classroom activities present rich and diverse characteristics. Excellent teachers need to keep pace with the times, strengthen education and maintenance, strive to create a good learning atmosphere for everyone, and enhance everyone's learning effect. The "SPOC + deep learning" hybrid model of college spoken English provides a reference for public classroom education. When designing online or offline classroom forms, you can fully consider the use of high-level thinking activities such as analysis, evaluation, and innovation to promote students' deep learning and further improve the effect of learning. Through the labeled data to train, the error is transmitted from top to bottom to fine-tune the network. Based on the parameters of each layer obtained in the first step, the parameters of the whole multilayer model are further optimized. This step is a supervised training process. The first step is similar to the random initialization initial value process of the neural network. Because the first step is not random initialization, but obtained by learning the structure of input data, this initial value is closer to the global optimization, so it can achieve better results. Therefore, the good effect of deep learning is largely due to the first step of feature learning.

Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work.

References

- [1] J. Simarmata, A. Djohar, J. Purba, and E. A. Juanda, "Design of a blended learning environment based on merrill's principles," *Journal of Physics: Conference Series*, vol. 954, no. 1, Article ID 012005, 2018.
- [2] C. Dziuban, C. R. Graham, P. D. Moskal, A. Norberg, and N. Sicilia, "Blended learning: the new normal and emerging technologies," *International Journal of Educational Technology in Higher Education*, vol. 15, no. 1, p. 3, 2018.
- [3] M. H. T. Kho, K. S. Chew, M. N. Azhar et al., "Implementing blended learning in emergency airway management training: a randomized controlled trial," *BMC Emergency Medicine*, vol. 18, no. 1, p. 1, 2018.
- [4] R. . Owston, "Empowering learners through blended learning," *International Journal on E-Learning*, vol. 17, no. 1, pp. 65–83, 2018.
- [5] N. Popovic, T. Popovic, I. Rovcanin Dragovic, and O. Cmiljanic, "A Moodle-based blended learning solution for physiology education in Montenegro: a case study," *Advances in Physiology Education*, vol. 42, no. 1, pp. 111–117, 2018.
- [6] M. D. R. C. Alcántar and J. C. Montes, "Percepción de los Estudiantes sobre el Aprendizaje Significativo y Estrategias de Enseñanza en el Blended Learning[J]," *Reice Revista Iberoamericana Sobre Calidad Eficacia Y Cambio En Educación*, vol. 16, p. 1, 2018.
- [7] Z. Fakhir and M. A. Ibrahim, "The effect of blended learning on private school students' achievement in English and their attitudes towards it," *English Language and Literature Studies*, vol. 8, no. 2, p. 39, 2018.
- [8] J. C. Almenara and V. Marín-Díaz, "Blended learning y realidad aumentada: experiencias de diseo docente Blended learning and augmented reality: experiences of educational design," *RIED. Revista Iberoamericana de Educación a Distancia*, vol. 21, no. 1, pp. 57–74, 2018.
- [9] M. P. Bergstr and V. Lindh, "Developing the role of Swedish advanced practice nurse (APN) through a blended learning master's program: consequences of knowledge organisation," *Nurse Education in Practice*, vol. 28, pp. 196–201, 2018.

- [10] R. Crawford, L. Jenkins, and L. Jenkins, "Making pedagogy tangible: developing skills and knowledge using a team teaching and blended learning approach," *Australian Journal of Teacher Education*, vol. 43, no. 1, pp. 127–142, 2018.
- [11] K. Matheos and M. Cleveland-Innes, "Blended learning: enabling higher education reform," *Revista Eletrônica de Educação*, vol. 12, no. 1, pp. 238–244, 2018.
- [12] H. Shu and X. Gu, "Determining the differences between online and face-to-face student-group interactions in a blended learning course," *The Internet and Higher Education*, vol. 39, pp. 13–21, 2018.

Research Article

Dense Dilated Attentive Network for Automatic Classification of Femur Trochanteric Fracture

Yuxiang Kang,¹ Jie Yu,² Zhipeng Ren,¹ Guokai Zhang ,³ Wen Cao,¹ Yinguang Zhang ,¹ and Qiang Dong ¹

¹Department of Orthopaedics, Tianjin Hospital, Tianjin 300211, China

²Qingdao Central Hospital, Qingdao 266042, China

³School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

Correspondence should be addressed to Yinguang Zhang; ygz_yyy@ieee.org and Qiang Dong; dongqiangtianjin@126.com

Received 12 August 2021; Accepted 11 September 2021; Published 22 October 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Yuxiang Kang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Automatic classification of femur trochanteric fracture is very valuable in clinical diagnosis practice. However, developing a high classification performance system is still challenging due to the various locations, shapes, and contextual information of the fracture regions. To tackle this challenge, we propose a novel dense dilated attentive (DDA) network for more accurate classification of 31A1/31A2/31A3 fractures from the X-ray images by incorporating a DDA layer. By exploiting this layer, the multiscale, contextual, and attentive features are encoded from different depths of the network and thus improving the feature learning ability of the classification network to gain a better classification performance. To validate the effectiveness of the DDA network, we conduct extensive experiments on the annotated femur trochanteric fracture data samples, and the experimental results demonstrate that the proposed DDA network could achieve competitive classification compared with other methods.

1. Introduction

Femur trochanteric fracture is one of the most commonly occurred fractures among elderly people. Especially, with the rapid growth of the aging population worldwide, the occurrence of this fracture increases rapidly which severely threatens the health of elderly people. Moreover, since this fracture could lead to high mortality rates and dramatically affect the quality of patients' life, effective and timely treatment is essential to relieve the pain of patients during the clinical diagnosis. Currently, the most effective way to diagnose this disease is by utilizing medical imaging such as X-rays or computed tomography (CT) to classify the types of fractures and then applying an appropriate treatment plan based on the corresponding diagnosis result. Typically, the OA/OTA classification criterion has been the most frequent and reliable used method to diagnose the condition of the fracture in the clinical fracture diagnosis. In this criterion, there are three types, e.g., 31A1, 31A2, and 31A3 (as shown in Figure 1), where 31A1 represents the simple pertrochanteric

fracture, 31A2 denotes the multifragmentary pertrochanteric fracture and lateral wall incompetent (≤ 20.5 mm) fracture, and 31A3 is the intertrochanteric (reverse obliquity) fracture [1]. Nevertheless, the conventional diagnosis method inspects patient images slice by slice which is usually tedious and time-consuming for the radiologists, and moreover, with different clinical experiences of radiologists, the final diagnosis result is liable to be empirical and subjective, which may hamper making the follow-up treatment plan. To tackle this challenge, a practicable way is to design a fracture computer-aided system [2–6] that helps the radiologist classify the fracture types automatically. In the past, a considerable number of researches have been proposed, for example, Demir et al. [7] developed a novel exemplar pyramid method for the humerus fracture; it extracted histograms of oriented gradients and local binary pattern features from the input images and then combined them with four conventional classifiers to classify the fractures. Boudissa et al. [8] explored the influence of semiautomatic bone-fragment segmentation on the reproducibility of the

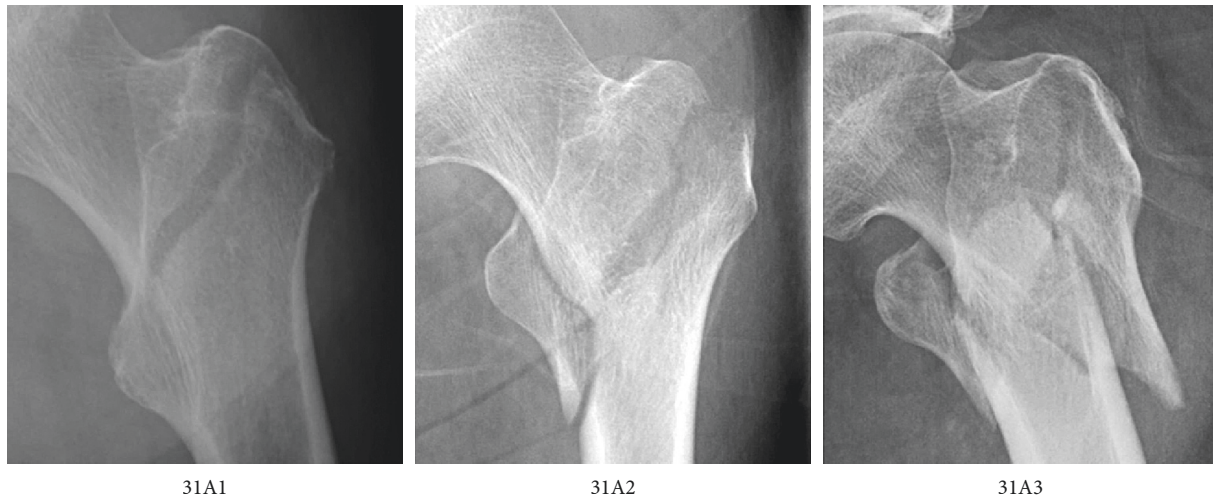


FIGURE 1: The fracture samples of 31A1, 31A2, and 31A3.

fracture classification, and it claimed that with the assistance of this technique, the classification accuracy of the fracture could be effectively improved. Additionally, in [9], the author proposed a 3D classification intertrochanteric fractures system which used the Hausdorff distance-based K-means method to classify the fractures into five types; the experimental results found that the unsupervised K-means method could gain promising classification performance with clinical significance. Cho et al. [10] evaluated the 3D CT images for boosting the diagnosis performance of femur trochanteric fractures, and it is shown that incorporating the CT could efficiently improve the reproducibility of stability for femur trochanteric fractures. Mall et al. [11] utilized different machine learning methods with the gray level cooccurrence matrix (GLCM) to classify the categories of fractures or no fractures; it proved that the proposed method could gain significant improvement on different evaluation metrics. Despite the great success, those methods have achieved in classifying the fracture task; those have deficiencies in capturing the robust and high-level semantic features due to hand-crafted feature predefinition.

Recently, the deep convolution neural network (CNN) has been proved its effectiveness in many computer vision tasks [12–16]. For instance, Lindsey et al. [17] suggested a deep CNN model based on the UNet structure achieving the automatic detection of the wrist fractures, and then it was evaluated on two different datasets; the result demonstrated that the proposed model could boost the clinical diagnosis performance. Then, Krogue et al. [18] labeled 3026 hip fractures and trained them with the DenseNet to achieve the automatic detection of hip fractures. Similarly, in [19], the authors utilized the faster R-CNN [20] to locate and classify the distal radius fractures automatically, and it obtained the mean average precision score of 0.866 at that time. To learn more high-level features, in [21], the authors employed a cropping process with the Inception V3 network to filter the unnecessary parts and thus leading to an improvement of the fracture detection. Besides, in [22], it used the Inception-ResNet faster R-CNN architecture to

construct a wrist fracture detection model and tested it on the unseen dataset, which proved that the designed model could gain high sensitivity and specificity.

Although those methods, especially the CNN ones, have gained promising results on the fracture classification task, an automatic fracture classification model should be simple and stable and provide effective information for the follow-up treatment plan. Specially, the femur trochanteric fractures usually have various locations, shapes, and contextual information in the clinical practice, which make it challenging to achieve a higher classification performance. Moreover, few works have considered the contextual information at different scales which may further limit the capability of the classification models. To tackle those challenges and efficiently improve the ability to learn strong representations from the fracture regions, in this paper, we develop a dense dilated attention (DDA) network to aggregate the multiscale, contextual, and attentive features from the femur trochanteric fracture region. Specially, in our DDA network, we incorporate the dense connection with dilated convolution by utilizing different dilated rates to learn the multiscale representations, and meanwhile, the dense connection could also alleviate the vanishing gradient problem and enable the network to reuse the hierarchical features. Furthermore, a dilated attention (DA) module is designed which encourages the network to encode more contextual and attentive representations automatically. To validate the effectiveness of DDA network, we perform extensive experiments on the femur trochanteric fracture images, and the experimental results show that our proposed DDA network could efficiently improve the classification performance by successfully extracting the discriminative features from the input image.

The rest of the paper is organized as follows: Section 2 presents the details of our proposed DDA network, and in Section 3, we first introduce the experimental data and evaluation metrics and then show the comparison results of different experiment settings. Finally, an elaborate discussion and conclusion of this paper are given in Section 4.

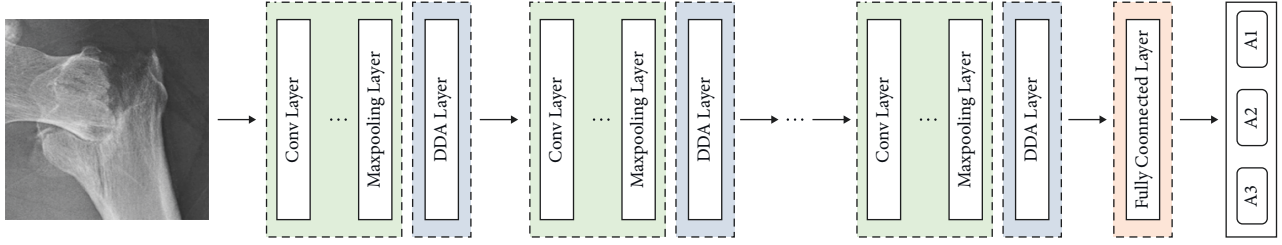


FIGURE 2: The main architecture of our proposed DDA network, where DDA layer is the dense dilated attention module.

2. Methodology

The automatic classification of femur trochanteric fracture is a challenging task due to its complex contextual information and various fracture regions. Hence, improving the network ability to extract multiscale representations, contextual information and intensity details are particularly important for accurate femur trochanteric fracture classification. To address those above challenges, a DDA network is developed for the accurate classification of the fracture categories; in the following subsections, we will provide the detailed descriptions of the network architecture and DDA module.

2.1. Network Architecture. As illustrated in Figure 2, given the X-ray images as the input of the DDA network, it first passes a series of convolution layers, max-pooling layers, and then a DDA module is implemented in the middle of the network to refine the feature representations, which will be described in detail in Subsection 2.3. After that, the final prediction category is output by a fully connected (FC) layer with the softmax activation in an end-to-end manner. The detailed parameters of the network are shown in Table 1. Notably, to preserve more spatial information of the image, we do not use the stride in the convolution layer. Specially, the ReLU activation is used to learn more nonlinear information, and batch normalization layer is utilized after each nonlinear activation to accelerate the convergence of the network.

2.2. Dilated Convolution. In our DDA module, we employ the dilated convolution to enlarge the receptive field without losing feature map resolution. Moreover, as the receptive field increases, it also provides more multiscale contextual features from the input [23]. Specifically, the dilated convolution could be divided into three steps: (1) sampling the input feature map based on the dilated rate; (2) conducting the convolution operation on the sampled values; (3) merging the obtained sampled values to a new feature map. Here, we denote the kernel size of the convolution layer as $k \times k$, and then the output feature map dimension of the traditional convolution layer I_c could be calculated as

$$I_c = \frac{W - k + 2p}{s} + 1, \quad (1)$$

TABLE 1: Parameters setting of the DDA network, where ‘‘Conv’’ denotes the convolution layer, and ‘‘FC’’ represents the fully connected layer.

Layer name	Output size	Filter size	Filter number
Conv1	512×512	3×3	256
Max-pooling	256×256	—	—
Conv2	256×256	3×3	128
Max-pooling	128×128	—	—
DDA layer	128×128	—	—
Conv3	128×128	3×3	64
Max-pooling	64×64	—	—
Conv4	64×64	3×3	64
Max-pooling	32×32	—	—
DDA layer	32×32	—	—
Conv5	32×32	3×3	32
Max-pooling	16×16	—	—
Conv6	16×16	3×3	32
Max-pooling	8×8	—	—
DDA layer	8×8	—	—
FC	Softmax	—	3

where W is the dimension of the input feature map; p and s are the padding size and stride, respectively. For the dilated convolution, its output dimension I_c^d could be defined as

$$I_c^d = \frac{W - k - r + 2p}{s} + 1. \quad (2)$$

Notably, when the stride is set as 1, the receptive field of the dilated convolution layer R_l could be formulated as

$$R_l = R_{l-1} + (k - 1) * r_l, \quad (3)$$

where r_l is the dilated rate of l th layer, and through this operation, the receptive field could increase rapidly. Finally, the dilated convolution layer could be given as

$$y[i] = \sum_m x[i + r \cdot m] \cdot w[m], \quad (4)$$

where $x[i]$, $y[i]$ are the input and output of the i th position, separately; $w[m]$ represents the learnable parameters of m th filter.

2.3. Dense Dilated Attention Layer. As illustrated in Figure 3, the aim of our dense dilated attention (DDA) layer is to learn more contextual and multiscale features across different layers to leverage the classification performance of the DDA network [24]. Note that the shallow layers usually contain the position information of the input, while the deep layers have high-level

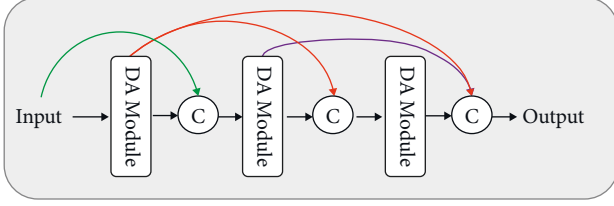


FIGURE 3: The structure of the dense dilated attention layer.

semantic representations. Therefore, combining those features across different layers could enhance the discrimination capability on fracture regions. Additionally, in order to guide the network's focus on the most salient regions from different receptive fields, a DA module is integrated into the DDA layer. Specially, taking the previous input feature map from the previous layer, it first passes through the DA module to learn the attentive and contextual features, and then the obtained ones are concatenated with the previous inputs as the input of the next layer. Note that the DDA layer mainly contains three DA modules with dilated rates of $\{1, 2, 3\}$, respectively, and its detailed structure is shown in Figure 4.

Mathematically, we denote the input of each DA module as $F_D^r \in \mathbb{R}^{W \times H \times C}$ with dilated rate of r , where $W \times H \times C$ represents the width, height, and channel numbers of F_D^r . Then, we adopt three 1×1 convolution layers to transform the F_D^r to three embeddings $\Phi \in \mathbb{R}^{W \times H \times \hat{C}}$, $\eta \in \mathbb{R}^{W \times H \times \hat{C}}$, $\tau \in \mathbb{R}^{W \times H \times \hat{C}}$, separately:

$$\phi = W_\phi(F_D^r), \quad \eta = W_\eta(F_D^r), \quad \tau = W_\tau(F_D^r), \quad (5)$$

where $W_\phi(\cdot)$, $W_\eta(\cdot)$, and $W_\tau(\cdot)$ denote the corresponding convolution operation; \hat{C} represents the channel number of those embeddings. After that, ϕ , η , and τ are flattened to the dimension of $\hat{C} \times HW$. To gain the contextual relation of F_D^r , a matrix multiplication between ϕ and η is applied, which can be given as

$$M = \phi^T \times \eta, \quad (6)$$

where $M \in \mathbb{R}^{HW \times HW}$ is the similarity matrix. Next, a softmax activation S is employed to normalize M to the interval of $[0, 1]$, and it could be formulated as

$$\tilde{M} = S(M). \quad (7)$$

Then, the attentive feature map $A \in \mathbb{R}^{HW \times HW}$ is gained by multiplying the \tilde{M} with τ , and it could be formulated as

$$A = \tilde{M} \times \tau^T. \quad (8)$$

Therefore, the final output of the DA module F_{DA}^r is defined as

$$F_{DA}^r = F_{DA}^r + W_\delta(A), \quad (9)$$

where $W_\delta(\cdot)$ is the 1×1 convolution operation. By adopting the hierarchical DA modules with the dense connection, the DDA layer could not only extract the multiscale features from different receptive fields but also learn the attentive and contextual information from the input.

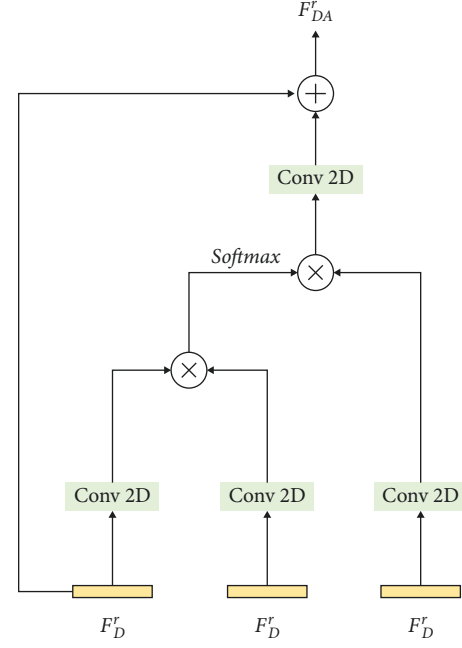


FIGURE 4: The structure of the DA module.

2.4. Training Loss Function. Denote the output feature map from the fully connected layer as $\{x_1, x_2, x_3\}$ and the corresponding label as $\{l_1, l_2, l_3\}$. To gain the predicted scores of each class, we apply a softmax activation function $\sigma(\cdot)$, which could be given as

$$\sigma(x)_i = \frac{e^{x_i}}{\sum_{n=1}^3 e^{x_n}}, \quad \text{for } i = \{1, 2, 3\}. \quad (10)$$

Additionally, to optimize the network, we use the binary cross-entropy \mathcal{L}_{bce} as the loss function, which could be formulated as

$$\mathcal{L}_{\text{bce}} = \sum_{j=1}^N g_j \log p_j + \sum_{j=1}^N (1 - g_j) \log(1 - p_j), \quad (11)$$

where N is the number of the data samples, and $p_j \in [0, 1]$, $g_j \in \{l_1, l_2, l_3\}$ are the predicted probability and corresponding true label.

3. Experiment

In this section, we conduct extensive experiments to validate the effectiveness of the DDA network. Specially, we first introduce the experimental data, implementation details, and evaluation metrics. Then, we compare the experiments with different amounts of data samples. Next, an ablation analysis of the DDA layer and DA module is explored to validate their efficiencies for this classification task. Finally, we reimplement some other fracture classification methods to compare them with our proposed DDA network.

3.1. Dataset. The total number of the experimental dataset is 390, and it consists of three categories 31A1, 31A2, and 31A3 with the amount of 117, 125, and 128, respectively. The mean

age of the patients is 65, the maximum age is 91, and the minimum age is 26. All the categories of the experimental data are annotated by three traumatic orthopedic specialists with more than 15 years of experience. Notably, the final category of data is based on the AO/OTA criterion. Since the initial resolution of the image is 1417×1772 , we crop the region of interest (ROI) with the maximum bounding-box, and then we resize those ROIs to 512×512 before input them into the DDA network to accelerate the training process of the network.

3.2. Implementation Details. In our experiments, since the initial resolution of the image is large, we first resize the image input to 512×512 . Moreover, to alleviate the overfitting problem, data augmentation is utilized to generate more data samples, and it includes random rotation, flip, and contrast. The whole network is implemented by the PyTorch deep learning library, and it is optimized by Adam. The initial learning rate is set as 0.001 and it decreases by 0.1 after 10 epochs. To accelerate the training process, we use the NVIDIA GeForce RTX 2070 Graphics Card, and the batch-size is set as 2.

3.3. Evaluation Metrics. To evaluate the performance of the proposed DDA network, we apply four evaluation metrics; here, we denote the true positive, false positive, true negative, and false negative as TP, FP, TN, and FN. Then, the accuracy which calculates the correct prediction among the total numbers of samples could be calculated as

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (12)$$

The sensitivity measures the ratio of correct TP prediction to the whole number of true positive samples. It could be formulated as

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (13)$$

The specificity calculates the ratio of correct TN to the whole numbers of false positive samples, and it is given as

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (14)$$

The receiver operating characteristic (ROC) curve is the most used graphical plot to measure the performance of the classifier, and the area under curve (AUC) is the score to measure the classification performance in which the higher score indicates the better distinguishing performance.

3.4. Data Sample Analysis of DDA Network. In this section, we first explore the influence of different data samples on the classification performance of the DDA network. Here, we divide our training data to 20%, 40%, 60%, 80%, and 100%, while the amount of the testing data samples is unchanged. The comparison result is shown in Figure 5, and it can be concluded that the more the data samples, the better performance the network would gain.

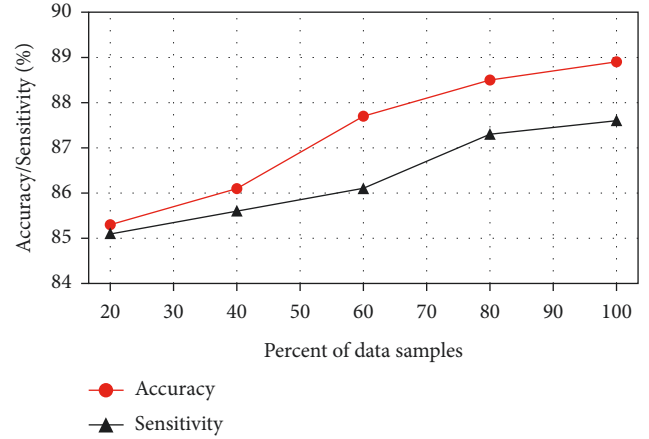


FIGURE 5: The accuracy/sensitivity of different percent of data samples.

That could be suggested that with more data samples, the network can extract the image features more sufficiently and effectively.

3.5. The Effectiveness of DDA Layer. The aim of our DDA layer is to learn the multiscale, attentive, and contextual information from the input image. Therefore, in this section, we conduct experiments w/o the DDA layer to explore its effects on the final classification performance. Moreover, since different depths of layers contain discriminative representations, therefore we also test the effectiveness of DDA layer on different depth locations. As reported in Table 2, DDA-1, DDA-2, and DDA-3 represent the depth location of locating DDA layer, in which the smaller value denotes the shallower depth location of the DDA network, and the DDA-W and DDA-O denote the network with or without the DDA layer. From the result, we observe that the best performance is achieved by the “DDA-W,” which can be explained that with the proposed DDA layer, the network is able to learn more high-level representations and then boost the classification performance.

3.6. Impact of DA Module. Different from the conventional dense connection, we develop a DA module that can guide the network to capture more attentive information with self-attention from different receptive fields. To validate the effectiveness of the proposed DA module, we compare three different network settings: with DA module; without DA module; with DA module (dilated rate as 1); with DA module (dilated rate as 2); with DA module (dilated rate as 3). The comparison result is shown in Table 3; the experimental result demonstrates that adopting the designed DA module could efficiently improve the classification performance compared with the setting without DA module. Moreover, with different dilate rates, the network tends to gain different performance; however, the best one is achieved by combining those three dilate rate settings, with the accuracy, sensitivity, specificity, and AUC of 88.9%, 87.6%, 85.9%, and 0.97, respectively.

TABLE 2: The effectiveness of DDA layer.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
DDA-1	88.0	86.0	84.3	0.91
DDA-2	88.1	86.3	85.1	0.92
DDA-3	88.5	87.1	85.3	0.94
DDA-O	87.3	85.3	82.1	0.85
DDA-W	88.9	87.6	85.9	0.97

TABLE 3: The comparison results of different DA settings, where “DA-W” and “DA-O” denote the network w/o DA module, and “DA($r=1,2,3$)” represents the DA module with dilate rates of 1, 2, and 3, respectively.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
DA ($r=1$)	88.2	87.2	85.1	0.93
DA ($r=2$)	88.1	86.8	84.7	0.91
DA ($r=3$)	88.3	87.0	85.2	0.92
DA-O	87.4	85.9	82.7	0.88
DA-W	88.9	87.6	85.9	0.97

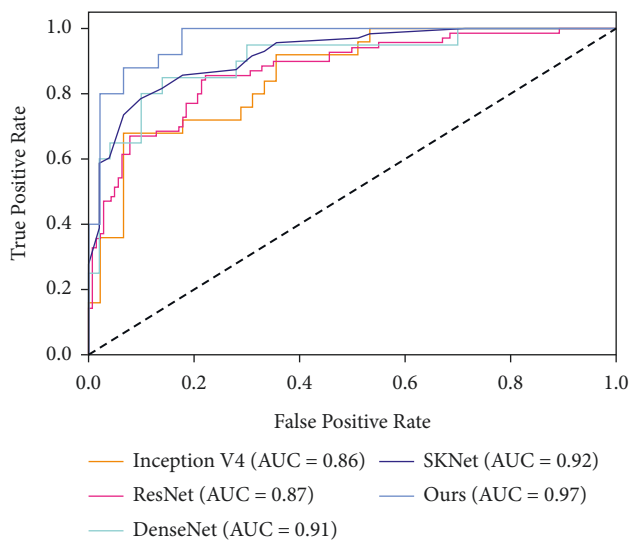


FIGURE 6: The AUC scores of different methods.

3.7. Comparison with Other Methods. To further evaluate the performance of the DDA network, in this section, we compare the results with different classification methods. As shown in Figure 6, we first compare our method with some baseline classification methods: Inception V4 [25], ResNet [26], DenseNet [27], and SKNet [28]. Note that we reimplement those methods and all the parameter settings are based on the default values. From the result, we observe that our proposed network could gain the highest AUC score of 0.97, which proves the effectiveness of our proposed network. Furthermore, we also report the comparison result with some other fracture methods; despite some of them are not with the same classification task, we reimplement those methods on the same dataset. As shown in Table 4, our method gains the best performance among all the evaluation metrics.

TABLE 4: The comparison results with other methods.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Exemplar [7]	87.5	85.7	82.0	0.80
K-means [9]	87.3	86.1	82.3	0.79
GLCM [11]	87.2	85.9	82.5	0.83
Inception V4 [25]	87.8	86.1	82.3	0.86
ResNet [26]	87.5	86.7	81.9	0.87
DenseNet [27]	87.2	86.3	83.4	0.91
SKNet [28]	87.1	86.8	83.6	0.92
Ours	88.9	87.6	85.9	0.97

4. Conclusion

In this paper, a DDA network is designed to achieve the classification of femur trochanteric fracture from X-ray images automatically. Since the fracture usually comes with various locations, shapes, and contextual information in the clinical practice, a novel DDA layer is developed which can automatically extract the multiscale, contextual, and attentive features to enhance the feature learning ability for achieving a more accurate classification performance. Extensive experiments on the annotated femur trochanteric fracture demonstrate that the proposed DDA network could gain competitive performance on this classification task. In future work, we will extend our work to different fracture classification tasks and collect more data samples to make the model more robust.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was sponsored by the Science and Technology Talent Cultivation Project of Tianjin Health Commission (KJ20215) and Research Project of Tianjin Sports Bureau (21DY014).

References

- [1] E. Meinberg, J. Agel, C. Roberts, M. Karam, and J. Kellam, “Fracture and dislocation classification compendium-2018,” *Journal of Orthopaedic Trauma*, vol. 32, no. 1, pp. S1–S10, 2018.
- [2] W. Li, Z. Wang, L. Zhang, L. Qiao, and D. Shen, “Remodeling pearson’s correlation for functional brain network estimation and autism spectrum disorder identification,” *Frontiers in Neuroinformatics*, vol. 11, p. 55, 2017.
- [3] X. Xu, W. Li, M. Tao et al., “Effective and accurate diagnosis of subjective cognitive decline based on functional connection

- and graph theory view,” *Frontiers in Neuroscience*, vol. 14, 2020.
- [4] P. Shan, F. Yang, H. Qi et al., “Alteration of MDM2 by the small molecule YF438 exerts antitumor effects in triple-negative breast c,” *Cancer Research*, vol. 81, no. 15, pp. 4027–4040, 2021.
 - [5] D. Gong, L. Wu, J. Zhang et al., “Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study,” *The Lancet Gastroenterology & Hepatology*, vol. 5, no. 4, pp. 352–361, 2020.
 - [6] Mo Zhang, B. Dong, and Q. Li, “Deep active contour network for medical image segmentation,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, Lima, Peru, October 2020.
 - [7] S. Demir, S. Key, T. Tuncer, and S. Dogan, “An exemplar pyramid feature extraction based humerus fracture classification method,” *Medical Hypotheses*, vol. 140, Article ID 109663, 2020.
 - [8] M. Boudissa, B. Orfeuvre, M. Chabanas, and J. Tonetti, “Does semi-automatic bone-fragment segmentation improve the reproducibility of the Letournel acetabular fracture classification?” *Orthopaedics and Traumatology: Surgery & Research*, vol. 103, no. 5, pp. 633–638, 2017.
 - [9] J. Li, S. Tang, H. Zhang et al., “Clustering of morphological fracture lines for identifying intertrochanteric fracture classification with Hausdorff distance-based K-means approach,” *Injury*, vol. 50, no. 4, pp. 939–949, 2019.
 - [10] Y.-C. Cho, P.-Y. Lee, C.-H. Lee, C.-H. Chen, and Y.-M. Lin, “Three-dimensional CT improves the reproducibility of stability evaluation for intertrochanteric fractures,” *Orthopaedic Surgery*, vol. 10, no. 3, pp. 212–217, 2018.
 - [11] 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*, pp. 1–6, IEEE, Allahabad, India, December 2019.
 - [12] J. Wang, Y. Yang, J. Mao, Z. Huang, C. Huang, and W. Xu, “Cnn-rnn: a unified framework for multi-label image classification,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, IEEE, Las Vegas, NV, USA, June 2016.
 - [13] H. Lee and H. Kwon, “Going deeper with contextual CNN for hyperspectral image classification,” *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4843–4855, 2017.
 - [14] Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. Lv, “Automatically designing CNN architectures using the genetic algorithm for image classification,” *IEEE transactions on cybernetics*, vol. 50, no. 9, pp. 3840–3854, 2020.
 - [15] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, “Medical image classification with convolutional neural network,” in *Proceedings of the 2014 13th international conference on control automation robotics vision (ICARCV)*, IEEE, Singapore, December 2014.
 - [16] S. S. Yadav and S. M. Jadhav, “Deep convolutional neural network based medical image classification for disease diagnosis,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–18, 2019.
 - [17] R. Lindsey, A. Daluiski, S. Chopra et al., “Deep neural network improves fracture detection by clinicians,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 45, 11591 pages, 2018.
 - [18] J. D. Krogue, K. V. Cheng, K. M. Hwang et al., “Automatic hip fracture identification and functional subclassification with deep learning,” *Radiology: Artificial Intelligence*, vol. 2, no. 2, Article ID e190023, 2020.
 - [19] E. Yahalomi, M. Chernofsky, and M. Werman, “Detection of distal radius fractures trained by a small set of X-ray images and Faster R-CNN,” *Advances in Intelligent Systems and Computing*, in *Proceedings of the Intelligent Computing-Proceedings of the Computing Conference*, pp. 971–981, Springer, Cham, London, UK, July 2019.
 - [20] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2016.
 - [21] M. D. V. Thurston, T. Mackinnon, and D. H. Kim, “Fracture detection with artificial intelligence: improved accuracy with region of interest focusing,” *European Congress of Radiology-ECR*, vol. 2018, 2018.
 - [22] Y. L. Thian, Y. Li, P. Jagmohan, D. Sia, V. E. Y. Chan, and R. T. Tan, “Convolutional neural networks for automated fracture detection and localization on wrist radiographs,” *Radiology: Artificial Intelligence*, vol. 1, no. 1, Article ID e180001, 2019.
 - [23] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2017.
 - [24] X. Wang, R. Girshick, and A. Gupta, “Non-local neural networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7794–7803, Salt Lake City, UT, USA, June 2018.
 - [25] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *Proceedings of the Thirty-first AAAI conference on artificial intelligence*, San Francisco CA USA, February 2017.
 - [26] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA, June 2016.
 - [27] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, HI, USA, July 2017.
 - [28] X. Li, W. Wang, X. Hu, and J. Yang, “Selective kernel networks,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 510–519, Long Beach, CA, USA, June 2019.

Research Article

AGVs Route Planning Based on Region-Segmentation Dynamic Programming in Smart Road Network Systems

Zheng Zhang, Juan Chen , and Qing Guo

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

Correspondence should be addressed to Juan Chen; jchen@mail.buct.edu.cn

Received 10 August 2021; Accepted 15 September 2021; Published 22 October 2021

Academic Editor: Wenbing Zhao

Copyright © 2021 Zheng Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a route-planning approach is proposed based on the region-segmentation Dynamic Programming (DP) algorithm for Automated Guided Vehicles (AGVs) in large Smart Road Network Systems (SRNSs) to deal with the problem of low route computation efficiency of the classical DP algorithm. We introduced an improved Markov Decision Process (MDP) to describe SRNSs, in which the SRNSs are divided into several regions according to the AGVs' start nodes and their goal nodes to improve the route-planning efficiency. Moreover, the route with the minimum number of turns is selected to reduce the system running time and energy cost in the following way: first, all the equidistant shortest routes are acquired from the AGVs' start nodes to their goal nodes using the improved DP algorithm; then, the routes are screened by calculating the angular deviation between all feasible routes and AGVs' initial directions, and the route with the fewest number of turns is taken as the shortest-time route. The simulation results verified that the proposed method can effectively solve the route-planning problem of AGVs in current SRNSs.

1. Introduction

Smart Road Network Systems (SRNSs) are important parts of the Smart World. Automated Storage and Retrieval Systems (AS/RSs) and Container Terminal Systems are typical Industry 4.0 application scenarios. Automated Guided Vehicles (AGVs) are the main tools, which enable automatic access to goods transportation without human labor. With the development of Smart City and Internet of Things technology, the route-planning algorithms of AGVs in SRNSs are commonly used not only in logistics industries and smart factories but also in intelligent transportation systems [1], energy transmission systems [2], and even network planning [3]. With the wide application of AGVs route planning, the goal of route planning is not only to find the route with the shortest distance from AGVs' start nodes to their goal nodes but also to minimize the system operating time and reduce energy cost. Moreover, the speed of the route-planning algorithm is also important. Therefore, AGVs route-planning problem in SRNSs has been widely studied.

The AGVs route planning in SRNSs belongs to the single-source shortest routing problem, and it has

similarities with trajectory optimization. They both can be solved using graph search algorithms [4, 5], sampling-based methods [6], and intelligent algorithms [7]. Dijkstra's algorithm is a classical route search algorithm, which has a simple structure and is robust and easy to implement, and it can meet requirements in practical scenarios. However, Dijkstra's algorithm belongs to Breadth-First Search (BFS), and thus it is not suitable for applying in scenarios with high real-time requirements [8–10]. In addition, when there are multiple equidistant shortest routes between the start node and the goal node, Dijkstra's algorithm can find only one of them. Even though there is improved Dijkstra's algorithm [11] that can obtain all equidistant shortest routes through iterations, the time complexity of Dijkstra's algorithm is $O(n^2)$, where n indicates the number of nodes in SRNSs. Therefore, it is not suitable for route planning in large SRNSs. AGVs interact with SRNSs all the time and an AGV has state s_t at each time t . An AGV takes action a on state s_t , and then it transforms to a new state s_{t+1} by taking reward r_{t+1} , and a sequence (i.e., $s_{t-1}, a_{t-1}, r_t, s_t, a_t, r_{t+1}, \dots$) is built by repeating the operation. The process of computing the shortest routes for AGVs is called the sequence decision process, and Markov Decision Process (MDP) is a typically

formulaic method for it. SRNSs described in this paper consist of several regularly arranged nodes, and the workflow can be described using MDP. However, we cannot completely copy the classical MDP to model the environment of the characteristics of different SRNSs.

Dynamic Programming (DP) algorithm is a branch of operations research, which was proposed by Bellman et al. [12] in the 1950s. DP algorithm is a method for solving optimization problems of the multistep decision process. DP can transform multistage problems into a series of single-stage problems [13], and it is suitable for solving optimization problems in large-scale environments. DP algorithm is the basic of Reinforcement Learning (RL) [14], and RL is an unsupervised algorithm based on the principles of reward and punishment. However, it is suitable for solving model-free problems, an AGV does not know which state has the bigger reward, and it needs to find the optimal policy to the goal state through trial and error [15]. Moreover, the classical DP algorithm iterates the sample set by randomly selecting starting point until the value function approaches the optimal policy. With the expansion of the scales of route-planning environments, the training time can become larger, and the computation amount of the DP algorithm will increase exponentially with the increasing number of nodes [16].

With the development of new communication technologies (e.g., 5G) and the increase of nodes' number in SRNSs, the real-time requirements for AGVs route-planning algorithm are becoming higher, and the classical DP algorithm is not suitable for the current development of Smart World. We propose a route-planning approach based on the region-segmentation DP algorithm for AGVs in SRNSs to address the low-efficiency problem of the classical DP algorithm. Because the classical MDP fails to accurately model the current SRNSs, we propose an improved MDP model for SRNSs. Since the number of samples is one of the important factors that affect the efficiency of the algorithm, a region-segmentation-based DP algorithm is proposed. Firstly, SRNSs are divided into several regions according to AGVs' start nodes and their goal nodes, and the objective is to reduce the number of training samples. Then, each node is assigned a value, and the DP algorithm is used to compute the value function of each region, and the objective is to find all equidistant shortest routes from the AGVs' start nodes to their goal nodes. Because AGVs can take more time and energy to turn, we introduce the following steps to determine routes with the least turns: (a) screening routes by calculating the angle deviation between all candidate routes and an AGV's initial direction and (b) choosing a route among remaining routes with the fewest number of turns and considering it the optimal route for the AGV. Moreover, we design different route-planning strategies according to whether an AGV is loaded or not, which improves the flexibility and efficiency of SRNSs, and the more nodes in the SRNSs, the more efficient our approach.

2. Related Works

The AGVs route-planning problem can be solved using Dijkstra's algorithm [17], A* algorithm [18], D* algorithm [19], Probabilistic Roadmap (PRM) [20], Rapid Random

Tree (RRT) [21], Artificial Potential Field (APF) [22], neural network [23], genetic algorithm [24], and ant colony algorithm [25]. Dijkstra's algorithm is a classical shortest route search algorithm, and it is simple and stable for performance. Guo et al. [11] improved Dijkstra's algorithm and found the shortest routes with the least distance and time consumption for AGVs in AS/RSSs. Li et al. [26] assumed that all obstacles were convex and found the shortest routes for AGVs without collision. However, the above achievements were all obtained in small-scale environments without considering the case of large-scale SRNSs (i.e., a large number of nodes and large size of the space). With the rapid development of SRNSs, the calculation amount of AGVs route planning will increase exponentially. But classical Dijkstra's algorithm does not consider the scenarios with a large number of nodes, and the system may be locked due to too much computation.

Yuan et al. [27] proposed Approximate Dynamic Programming (ADP), and they applied it to crane scheduling to reduce the operation time of the AS/RSSs. Novoa et al. [28] focused on Vehicle Routing Problem (VRP) from the perspective of ADP, and they proved that a DP-based algorithm could be applied to vehicle routing planning. Cimen et al. [29] proposed a vehicle routing optimization algorithm by combining ADP and MDP. Bahlawan et al. [30] introduced a method of system operation management and energy optimization based on the DP algorithm to minimize the energy cost of factories. Horiguchi et al. [31] focused on the routing problem of network packets, and they redefined the energy function by combining the DP algorithm and neural network. They proposed a method to find the optimal balance between queues' length and the routes' distance. Although many achievements have proved that the DP algorithm can be used to solve optimization problems in various situations, the classical DP algorithm needs to iterate the values of each node in the whole space many times until a stable and optimal policy is obtained. To reduce the computing time of DP, Ulmer et al. [32] proposed an offline value function prediction method, which introduced MDP into state space and combined action space with reward information. Desai et al. [33] proposed a preprocessing method of a random dynamic network according to vehicles' starting time and location. But they failed to fundamentally solve the problem of longer DP algorithm iteration time.

3. Rasterizing the SRNS as the Model

There are a large number of nodes arranged according to the regular form of rows and columns in SRNSs and nodes and the lines between nodes (i.e., the driving paths of AGVs) are typically presented as a regular rectangle. Therefore, SRNSs can be rasterized, and the grids' locations can be stored with spatial coordinates. Figures 1(a)–1(c) are $m \times n$ SRNSs constructed using grid methods [34], where m and n represent the numbers of rows and columns of the SRNSs, respectively. We establish the Cartesian coordinate system of the SRNSs: the upper-left corner is selected as the origin ($x=0, y=0$), the lower-

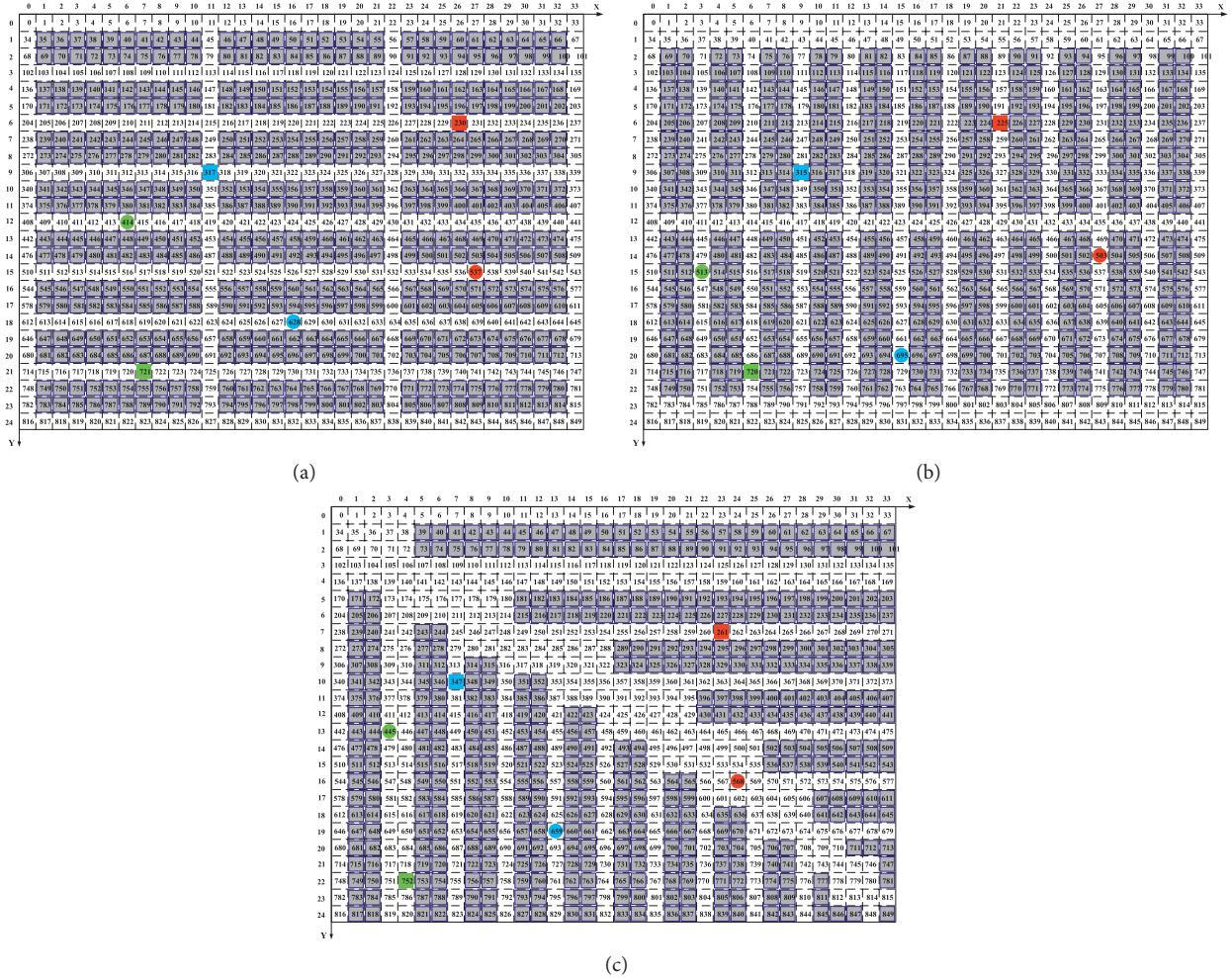


FIGURE 1: Rasterize large SRNSs as a model. (a) Horizontal row SRNSs; (b) vertical column SRNSs; (c) fish-bone SRNSs.

right corner defines the workspace size ($x = n - 1$, $y = m - 1$), and each node corresponds to a coordinate (a, b) ($a \leq n - 1$, $b \leq m - 1$). Grey grids represent obstacle nodes, white grids indicate the nodes that are available for AGVs, and numbers in grids indicate the number of nodes (i.e., the number of nodes starting from $(x = 0, y = 0)$ and marked as 1, 2, 3, . . . , n). Red, green, and blue solid squares represent the start nodes of AGVs, and solid circles of the corresponding color represent their goal nodes. The distance between every two adjacent nodes is equal and the SRNS is bidirectional (i.e., AGVs can travel bidirectionally), and each node can only be occupied by one AGV. The size of every obstacle is the same in SRNSs and the number of grids AGVs passed represents the travel distance for AGVs.

Figures 1(a) and 1(b) are commonly used SRNSs (e.g., AS/RSs), in which the shelves are arranged according to the rules of row and column, and there is a channel on both sides of each row (or column) of shelf for AGVs access. Figure 1(c) shows recently presented SRNSs due to the rapid development of the Smart World. In these fish-bone SRNSs, shelves are arranged on both sides of the main diagonal of the SRNSs, since it is inconvenient for AGV

access and the signal may be blocked, and it can lead to system deadlock of SRNSs; this environment model has no practical application.

3.1. General Markov Decision Process Framework. Markov Decision Process (MDP) is a quintuple (i.e., $\langle S, A, P, R, \gamma \rangle$), where S is the finite states set, A is the actions set, and P is the state transition probability matrix. Given the current state s and the next state s' , the state transition probability (i.e., $P_{ss'} = P[S_{t+1} = s' | S_t = s]$) represents the transition probability from state s to state s' . R is the rewards set; γ is the discount factor, which denotes that the influence of the goal state on each state is gradually weakened with the increase of distance between a node and the goal node, $\gamma \in [0, 1]$. There are two kinds of rewards of R set: (1) immediate reward R_t , which denotes the expectation of rewards from state s to state s' , that is, $R_t = E[R_{t+1} | S_t = s]$, and (2) accumulated reward G_t , which denotes the weighted sum of each immediate reward from the start state to the goal state.

$$G_t = R_{t+1} + \gamma R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}. \quad (1)$$

3.2. *The Markov Decision Process in SRNSs.* State value function $V(s)$ represents the expectation of accumulated reward under state s :

$$V(s) = E[G_t | S_t = s]. \quad (2)$$

Bellman equation [12] is the core of all MDP-based algorithms. AGVs can obtain the best route from start nodes to goal nodes according to Bellman equation. The derivation process of the Bellman equation is as follows:

$$\begin{aligned} V(s) &= E[G_t | S_t = s] \\ &= E[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} \dots | S_t = s] \\ &= E[R_{t+1} + \gamma(R_{t+2} + \gamma R_{t+3} \dots) | S_t = s] \quad (3) \\ &= E[R_{t+1} + \gamma G_{t+1} | S_t = s] \\ &= E[R_{t+1} + \gamma V(S_{t+1}) | S_t = s]. \end{aligned}$$

In equation (3), the value of state s consists of (i) the expectation of immediate reward and (ii) the expectation of the state's value for the next moment. State-policy transition probability $P_{s,s'}^\pi = \sum_{a \in A} \pi(a|s) P_{s,s'}^a$, represents the probability of an AGV transforming from state s to state s' .

3.2.1. *Smart Road Network System-Reward Matrix (SRNS-RM).* AGVs can obtain different immediate rewards at different nodes, and the reward/penalty obtained by AGVs is defined as follows: (1) if AGVs enter grey solid grids (i.e., obstacle nodes), they will get penalty: $-r$, where r is a positive integer; (2) if AGVs enter white grids, they will get penalty: $-\varepsilon$, where ε is a positive integer and $\varepsilon \ll r$; (3) if AGVs reach the goal node, they can obtain reward $+r^2$. Given an $m \times n$ SRNS, where m indicates the number of rows, that is, the number of nodes in a column, and n indicates the number of columns, that is, the number of nodes in a row, the SRNS-RM is

$$\text{SRNS-RM} = \begin{bmatrix} r_{11} = -\varepsilon & r_{12} = -\varepsilon & r_{13} = -r & \dots & r_{1n} = -\varepsilon \\ r_{21} = -\varepsilon & r_{22} = -r & r_{23} = -r & \dots & r_{2n} = -\varepsilon \\ r_{31} = -\varepsilon & r_{32} = -\varepsilon & r_{33} = -r & \dots & r_{3n} = -\varepsilon \\ \dots & \dots & \dots & \dots & \dots \\ r_{m1} = -\varepsilon & r_{m2} = -\varepsilon & r_{m3} = -\varepsilon & \dots & r_{mn} = +r^2 \end{bmatrix}. \quad (4)$$

In equation (4), the dimension of SRNS-RM is $m \times n$, which is the same as the SRNS. The elements (i.e., r_{ij}) in SRNS-RM ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$) indicate the rewards of an AGV reaching nodes in row i and column j . Consider that there are four obstacle nodes in the SRNS, that is, $\text{SRNS-RM}_{13} = -r$, $\text{SRNS-RM}_{22} = -r$, $\text{SRNS-RM}_{23} = -r$, and $\text{SRNS-RM}_{33} = -r$.

3.2.2. *Smart Road Network System-Transition Probability Matrix (SRNS-TPM).* The adjacencies of nodes in SRNSs are as follows: (1) there are no upper adjacent nodes for the nodes on the upper edge but only left, lower, and right adjacent nodes; (2) there are no left adjacent nodes for the

node on the left edge but only top, right, and bottom adjacent nodes; (3) there are no lower adjacent nodes for the node on the lower edge but only upper, left, and right adjacent nodes; (4) there are no right adjacent nodes for the node on the right edge but only top, left, and bottom adjacent nodes; (5) there are four adjacent nodes except for the nodes on edges. In principle, the probabilities are the same for an AGV entering any adjacent node in an $m \times n$ SRNS, that is, $P_{ij} = P[S_{t+1} = j | S_t = i] = \delta$, where $0 \leq \delta \leq 1$. The SRNS-TPM is

$$\text{SRNS-TPM} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \dots & P_{1n} \\ P_{21} & P_{22} & P_{23} & \dots & P_{2n} \\ P_{31} & P_{32} & P_{33} & \dots & P_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ P_{m1} & P_{m2} & P_{m3} & \dots & P_{mn} \end{bmatrix}. \quad (5)$$

In equation (5), P_{ij} indicates each element of SRNS-TPM, where i indicates AGV's current state (i.e., S_t), and j indicates its next state (i.e., S_{t+1}). The dimension of SRNS-TPM is $m \times n$, which is the same as the SRNS and $P_{11} = P_{12} = \dots = P_{mn} = \delta$; that is, the probability of AGV entering each adjacent node is the same.

3.2.3. *Criteria of the Optimal Policy Selection.* The objective of the classical MDP is to find the optimal policy that can reach the goal to maximize the accumulated reward G_t (see equation (2)). However, AGVs route planning in SRNSs is a multiobjective combinatorial optimization problem. On the one hand, the total routes' length should be minimized; on the other hand, total routes' length should be minimized, and AGVs' traveling time should be the least. Therefore, the criteria for choosing the optimal policy need to be modified.

When an AGV enters a node, it uploads coordinates of the current node to the server through 5G, Bluetooth, or other wireless technologies, so that the server can know the location of the AGV in SRNSs. We ignore the size of nodes in SRNSs, and consider that the distance between every two adjacent nodes is L_n ; the longitudinal length of the AGV is L_c ; and the traveling speed of the AGV is fixed as v . When the head of the AGV enters a node, the node is occupied, and the node is released when the tail of the AGV leaves. As shown in Figure 2, the time for the AGV to pass through every two adjacent nodes is $t_c = L_n + L_c/v$.

In many SRNSs (e.g., AS/RSs), an AGV can only travel in four directions, north, east, south, and west, and can only turn at nodes. It requires deceleration, stopping, autobigraphy, and acceleration, and the total time for turning is t_u . The route length from the start node to the goal node is shown in equation (6). Given that the starting time of an AGV is 0 s, the total time for the AGV traveling to the goal node is shown in equation (7).

$$L = (n_c - 1) \times L_n, \quad (6)$$

$$T = (n_c - 1) \times t_c + n_u \times t_u. \quad (7)$$

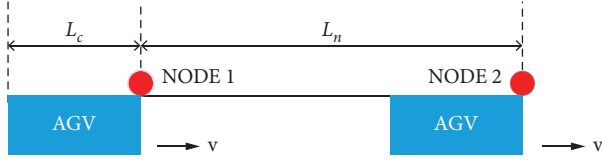


FIGURE 2: Time of AGVs passing through every two adjacent nodes.

In equations (6) and (7), n_c denotes the total number of nodes in the route and n_u denotes the number of turns in the route.

Given that the computing time of the server is t_s , the criteria of the optimal policy selection in SRNSs are as follows:

$$\min_{\pi} t_s, \quad (8)$$

$$\min_{\pi} L = (n_c - 1) \times L_n, \quad (9)$$

$$\min_{\pi} T = (n_c - 1) \times t_c + n_u \times t_u, \quad (10)$$

$$\text{s.t.} \quad \begin{cases} t_s > 0 \\ n_c > 1 \text{ and } L_n > 0 \\ 0 \leq n_u \leq n_p \\ t_u > 0 \end{cases} \quad n_p = 1, 2, \dots, n. \quad (11)$$

Equation (8) indicates that we should reduce route computing time of the server; equation (9) indicates that we should shorten routes' distance of AGVs; equation (10) indicates that we should minimize the total travel time of AGVs. In equation (11), n_p denotes the number of nodes in an AGV's route. Due to the fewer turns in AGVs' routes, with less energy cost of the SRNSs, reducing the number of turns in AGVs routes can reduce the energy cost.

3.3. AGVs Route-Planning Approach in SRNSs. The flow-chart of AGVs route-planning approach proposed in this paper is shown in Figure 3, which consists of four parts: (1) route-planning region segmentation; (2) computing all shortest routes by training value functions of the regions; (3) obtaining the feasible routes according to AGVs' status; and (4) selecting collision-free routes with the shortest length and the least travel time.

3.3.1. Route-Planning Region Segmentation. The most effective way to speed up training is to reduce the number of samples. We here proposed a region-segmentation-based DP algorithm, in which the SRNSs are divided into regions according to AGVs' start nodes and their goal nodes. The objective is to accelerate the convergence of the DP algorithm by excluding irrelevant nodes. Given the coordinates of an AGV's start node (X_s, Y_s) and the goal node (X_g, Y_g) , the left boundary of an SRNS X_l and the right boundary X_r , the upper left coordinates of the AGV's route-planning region $(X_{AGV_i}^{ul}, Y_{AGV_i}^{ul})$, the upper-right coordinates $(X_{AGV_i}^{ur}, Y_{AGV_i}^{ur})$, the coordinates of the lower-left corner

$(X_{AGV_i}^{dl}, Y_{AGV_i}^{dl})$, the coordinates of the lower right corner $(X_{AGV_i}^{dr}, Y_{AGV_i}^{dr})$, the length of the region along the x -direction L_{row} , and the length along the y -direction L_{col} , the AGVs' route-planning region is defined as follows:

(i) If $X_g - X_l < X_r - X_g$, that is, an AGV's goal node is closer to the left boundary of the SRNS, then

$$(X_{AGV_i}^{ul}, Y_{AGV_i}^{ul}) = (X_l, \min(Y_s, Y_g)), \quad (12)$$

$$(X_{AGV_i}^{ur}, Y_{AGV_i}^{ur}) = (\max(X_s, X_g), \min(Y_s, Y_g)), \quad (13)$$

$$(X_{AGV_i}^{dl}, Y_{AGV_i}^{dl}) = (X_l, \max(Y_s, Y_g)), \quad (14)$$

$$(X_{AGV_i}^{dr}, Y_{AGV_i}^{dr}) = (\max(X_s, X_g), \max(Y_s, Y_g)), \quad (15)$$

$$L_{row} = \max(X_s, X_g) - X_l, \quad (16)$$

$$L_{col} = |Y_g - Y_s|. \quad (17)$$

(ii) If $X_g - X_l > X_r - X_g$, that is, an AGV's goal node is closer to the right boundary, then

$$(X_{AGV_i}^{ul}, Y_{AGV_i}^{ul}) = (\min(X_s, X_g), \min(Y_s, Y_g)), \quad (18)$$

$$(X_{AGV_i}^{ur}, Y_{AGV_i}^{ur}) = (\min(X_s, X_g), \max(Y_s, Y_g)), \quad (19)$$

$$(X_{AGV_i}^{dl}, Y_{AGV_i}^{dl}) = (X_r, \min(Y_s, Y_g)), \quad (20)$$

$$(X_{AGV_i}^{dr}, Y_{AGV_i}^{dr}) = (X_r, \max(Y_s, Y_g)), \quad (21)$$

$$L_{row} = X_r - \max(X_s, X_g), \quad (22)$$

$$L_{col} = |Y_g - Y_s|. \quad (23)$$

The conventional region-segmentation methods only focus on locations of AGVs' start nodes and their goal nodes, but we take the relationship between them and the left (and right) boundaries of SRNSs into consideration. Our approach can reduce the number of samples without losing the accuracy of route planning. As shown in Figure 4, given that the start node of an AGV is node 43 ($x=3, y=5$) and the goal node is node 20 ($x=4, y=2$), the blue line indicates the feasible route from the start node to the goal node; the yellow shaded area represents the route-planning region obtained using the conventional region-segmentation methods, which does not include the feasible route, and the area in the red box indicates the region in our approach, which contains the feasible route.

3.3.2. Training Value Function of the Route-Planning Region. There are two kinds of commonly used DP algorithms: policy iteration-based DP algorithm and value iteration-based DP algorithm. The time of computing route using the DP algorithm based on policy iteration and value iteration in

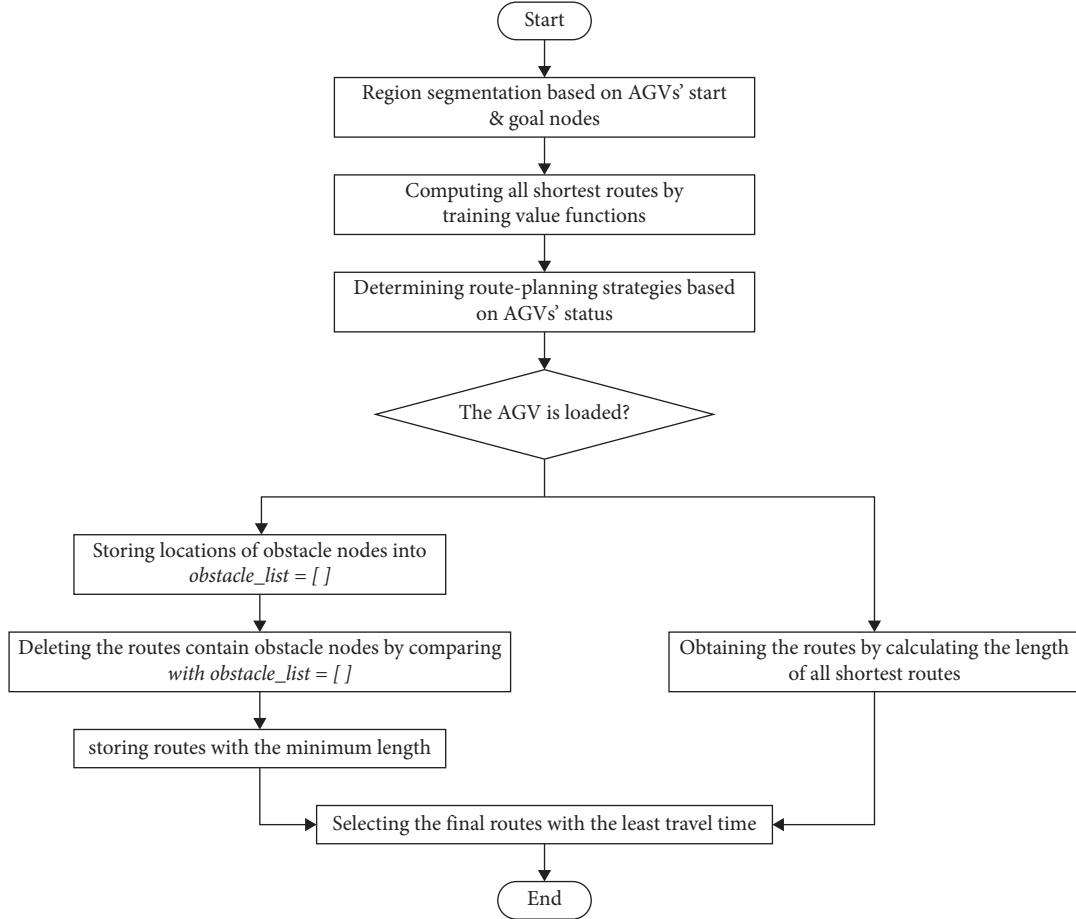


FIGURE 3: Flowchart of the AGVs route-planning approach proposed in this paper.

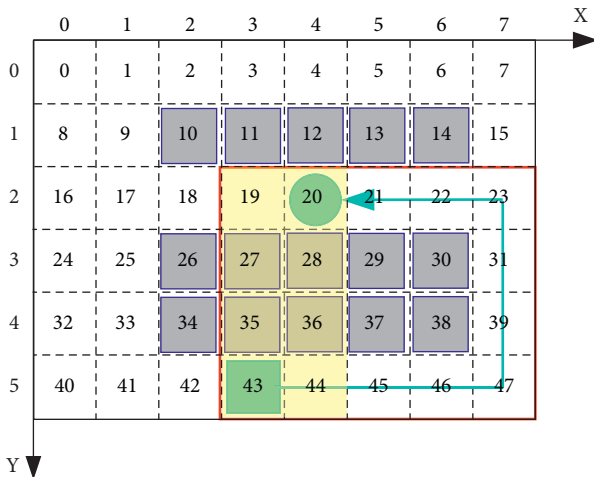


FIGURE 4: Conventional and improved region segmentation.

an $m \times n$ SRNS (see Figure 1(a)) is shown in Figure 5, where the blue line denotes the training time of the DP algorithm based on policy iteration, and the red line denotes that of the DP algorithm based on value iteration.

As can be seen from Figure 5, the training time of the DP algorithm based on value iteration is smaller than that of the DP algorithm based on policy iteration. In this scenario, the average time for route computing is improved by nearly

97.39%, and when the scales of SRNSs become larger, the difference of computing time between them becomes bigger. We employ the value iteration-based DP algorithm because it can adjust the policy while training rather than after the value function converges, which greatly reduces the number of iterations and significantly improves the efficiency of the algorithm.

The calculation of the optimal value function (i.e., $V^*(s)$) is a recursive process (refer to equation (3)). We need to evaluate the rewards of all adjacent states based on the current state, and the state with the largest reward is taken as the next state of the AGV. The method of computing the relationship between the current node and all adjacent nodes is

$$V^{T+1}(s) = \sum_{s+1} p(s_{t+1}|s_t, a_t) [r_{a_t}^{s_{t+1}} + \gamma \times V^T(s_{t+1})], \quad (24)$$

where $p(s_{t+1}|s_t, a_t)$ indicates the probability of taking action a_t to transfer to an adjacent state s_{t+1} based on the current state s_t ; $r_{a_t}^{s_{t+1}}$ is the reward of taking action a_t to transfer to state s_{t+1} ; γ is the discount factor, where $\gamma \in [0, 1]$; $V^T(s_{t+1})$ is the value of each state adjacent to state s_t .

Without considering obstacle nodes in SRNSs, the probability is the same for an AGV traveling to surrounding nodes; that is, $p(s_{t+1}|s_t, a_t) = 1/4$ and $\gamma = 1$. Equation (24) can be rewritten as

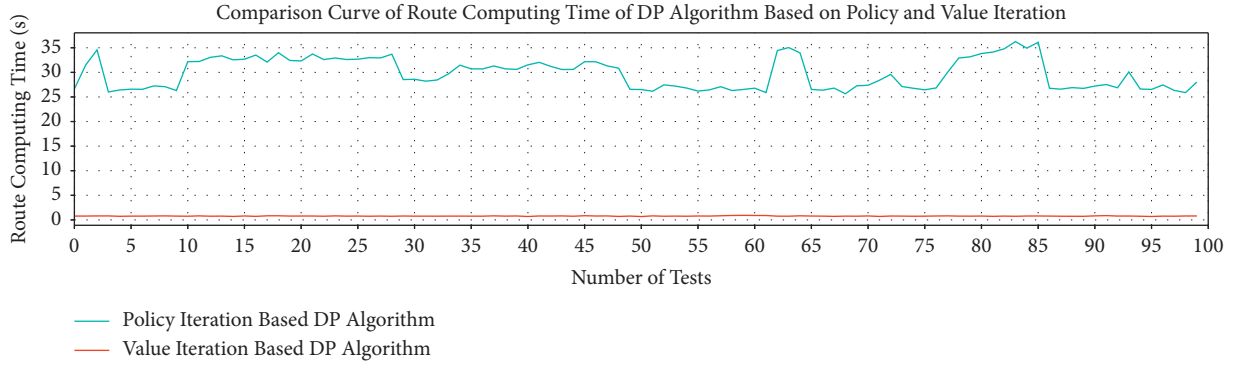


FIGURE 5: Curve of convergences between the DP algorithms based on policy iteration and value iteration.

$$V^{T+1} = \frac{1}{4} \times [(-1 + 1 \times V_u^T) + (-1 + 1 \times V_d^T) + (-1 + 1 \times V_l^T) + (-1 + 1 \times V_r^T)]. \quad (25)$$

In equation (25), V_u^T , V_d^T , V_l^T , and V_r^T denote the state values of upper, lower, left, and right adjacent nodes of AGV's current nodes, respectively.

3.3.3. Route-Planning Strategies Based on Status of AGVs. There are multiple obstacles on AGVs' routes, and the route-planning methods of AGVs in different status can be different. Taking AS/RSs as an example [35], as shown in Figure 6, the height of the bottom shelf from the ground is 75 cm, and the height of each no-load AGV is 50 cm. If an AGV needs to load, the tray on the AGV will be lifted to hold the bottom to lift the shelf. Because the height from the bottom shelf to the ground is greater than the height of a no-load AGV, a no-load AGV can pass under the shelf, whereas a loaded AGV (i.e., an AGV that is carrying a shelf) can not. So the loaded AGVs should avoid collision with not only other AGVs but also shelves placed above the nodes.

If AGVs are loaded, the route-planning method is as follows: (1) storing locations of obstacle nodes into a list $obstacle_list = []$ while initializing the environment; (2) repeating equation (25) to compute the value function of the route-planning regions and obtaining all feasible routes from AGVs' start nodes to their goal nodes; (3) comparing nodes in all feasible routes with nodes in $obstacle_list = []$ and deleting the routes that contain obstacle nodes; and (4) calculating the length of each collision-free route and storing routes with the minimum length. If AGVs are no-load ones, it is not necessary to establish $obstacle_list = []$ while initializing the environment, and we need not compare feasible routes to $obstacle_list = []$ but only to calculate the length of all feasible routes.

3.3.4. Obtaining Collision-Free Routes with the Shortest Length and the Least Travel Time. SRNSs are shaped like chessboards, and AGVs can travel in four directions: north, east, south, and west. Therefore, there may be more than one

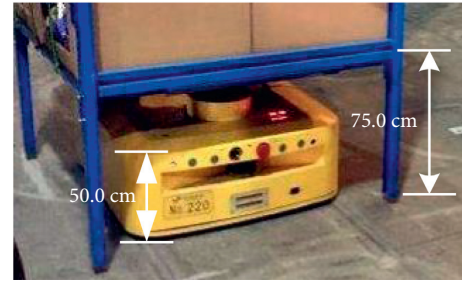


FIGURE 6: Height of bottom shelf and no-load AGV in AS/RSs.

shortest-distance route from AGVs' start nodes to their goal nodes. In some large SRNSs (e.g., AS/RSs), AGVs can only turn at nodes, and the time of the process becomes longer with the increase of AGVs' turning angles. The fewer angles, the smaller the time cost and the higher the system efficiency. We can find routes with the shortest travel time for AGVs through the following steps:

Step 1. Comparing the angles of each feasible route with AGVs' initial directions to determine their departure routes:

$$\theta_{init} = \min(|\theta_1^{start} - \theta_{AGV}^{start}|, |\theta_2^{start} - \theta_{AGV}^{start}|, |\theta_3^{start} - \theta_{AGV}^{start}|, |\theta_4^{start} - \theta_{AGV}^{start}|). \quad (26)$$

In equation (26), θ_{init} indicates the departure angle of AGVs from start nodes. We take the southward as 0° and start nodes as the origin; θ_1^{start} indicates the angle of route 1, θ_2^{start} indicates the angle of route 2, θ_3^{start} indicates the angle of route 3, and θ_4^{start} indicates the angle of route 4; θ_{AGV}^{start} indicates the angle of AGVs at start nodes.

Step 2. Computing the routes with the shortest travel time, that is, the routes with the fewest turns: traversing each node of feasible routes obtained in Step 1, deleting routes that contain nodes in $obstacle_list = []$, and then judging whether a node is a turn or not in remaining routes. The method of determining whether a node is a turn or not is as follows: considering the X and Y coordinates of upstream and downstream nodes, if a node satisfies equation (27), it is a turn:

$$|X_{i+1} - X_i| \neq |X_i - X_{i-1}| \text{ or } |Y_{i+1} - Y_i| \neq |Y_i - Y_{i-1}|. \quad (27)$$

In equation (27), i denotes the number of the current node, X_i is the X -coordinate of the current node, and Y_i is the Y -coordinate; $i + 1$ denotes the number of the downstream node, X_{i+1} is the X -coordinate, and Y_{i+1} is the Y -coordinate; $i - 1$ denotes the number of the upstream node, X_{i-1} is the X -coordinate, and Y_{i-1} is the Y -coordinate.

4. Simulation Case Studies

Simulation OS is *Windows 10, Intel Xeon W-2145 CPU @ 3.70 GHz x64-based processor*; programming software is *Python 3.7.3*; rasterized AS/RSs environment parameters are *25×34 Horizontal Row*, with 850 nodes in total, $L_c = 2.5 \text{ m}$ (note: the proposed algorithm is suitable for not only AS/RSs but also all SRNSs); $v = 0.5 \text{ m/s}$, and the times for AGVs traveling between two adjacent nodes (i.e., t_c) and turning at nodes (i.e., t_u) are both 5 s.

Case 1. Route planning for an AGV in AS/RSs from a designated start node to a designated goal node. The numbers of obstacle nodes of Figure 7 are as follows:

obstacle_list = [35, 36, 37, 38, 39, 40, 41, 42, 43, 44, . . . , 805, 806, 807, 808, 809, 810, 811, 812, 813, 814]. The start node of an AGV is node 56 ($x = 22, y = 1$), and the goal node is node 639 ($x = 27, y = 18$). The process of route planning is as follows.

4.1. Determining the Route-Planning Region for the AGV. Based on the start node and the goal node of every AGV and the AS/RS's boundary, we use equations (12)–(17) (or equations (18)–(23)) to obtain the region-segmentation result, as shown in the blue shaded part in Figure 7.

4.2. Modeling the AS/RS Using the Improved MDP. According to the definition in equations (16) and (17) (or equations (22) and (23)), we can obtain the following: $L_{row} = 12, L_{col} = 18, r = 10$, and $\varepsilon = 1$, and the probability of AGV traveling towards the node adjacent to the current node (i.e., state transition) $p = 1/4$. Based on equations (4) and (5), SRNS-RM and SRNS-TPM are as follows:

$$\begin{aligned}
 \text{PBS-RM} &= \begin{bmatrix} r_{1,1} = -1 & r_{1,2} = -10 & r_{1,3} = -10 & \cdots & r_{1,12} = -1 \\ r_{2,1} = -1 & r_{2,2} = -10 & r_{2,3} = -10 & \cdots & r_{2,12} = -1 \\ r_{3,1} = -1 & r_{3,2} = -1 & r_{3,3} = -1 & \cdots & r_{3,12} = -1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{18,1} = -1 & r_{18,2} = -1 & r_{18,3} = -1 & \cdots & r_{18,12} = +100 \end{bmatrix}, \text{PBS-TPM} \\
 &= \begin{bmatrix} P_{1,1} = \frac{1}{4} & P_{1,2} = \frac{1}{4} & P_{1,3} = \frac{1}{4} & \cdots & P_{1,12} = \frac{1}{4} \\ P_{2,1} = \frac{1}{4} & P_{2,2} = \frac{1}{4} & P_{2,2} = \frac{1}{4} & P_{2,3} = \frac{1}{4} & P_{2,12} = \frac{1}{4} \\ P_{3,1} = \frac{1}{4} & P_{3,2} = \frac{1}{4} & P_{3,2} = \frac{1}{4} & P_{3,3} = \frac{1}{4} & P_{3,12} = \frac{1}{4} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{18,1} = \frac{1}{4} & P_{18,2} = \frac{1}{4} & P_{18,2} = \frac{1}{4} & P_{18,3} = \frac{1}{4} & P_{18,12} = \frac{1}{4} \end{bmatrix}. \quad (28)
 \end{aligned}$$

4.3. Obtaining All the Equidistant Shortest Routes. By using the value iteration-based DP algorithm (see equation (25)), we can obtain all the feasible routes from AGV's start node to its goal node, as shown in Figures 8(a)–8(f).

4.4. Selecting a Route with the Least Travel Time. The route of an AGV in SRNS can be presented as $S(T)$, where S indicates the set of nodes' numbers in AGV's route, and T indicates the set of time when the AGV passes the nodes. We can choose

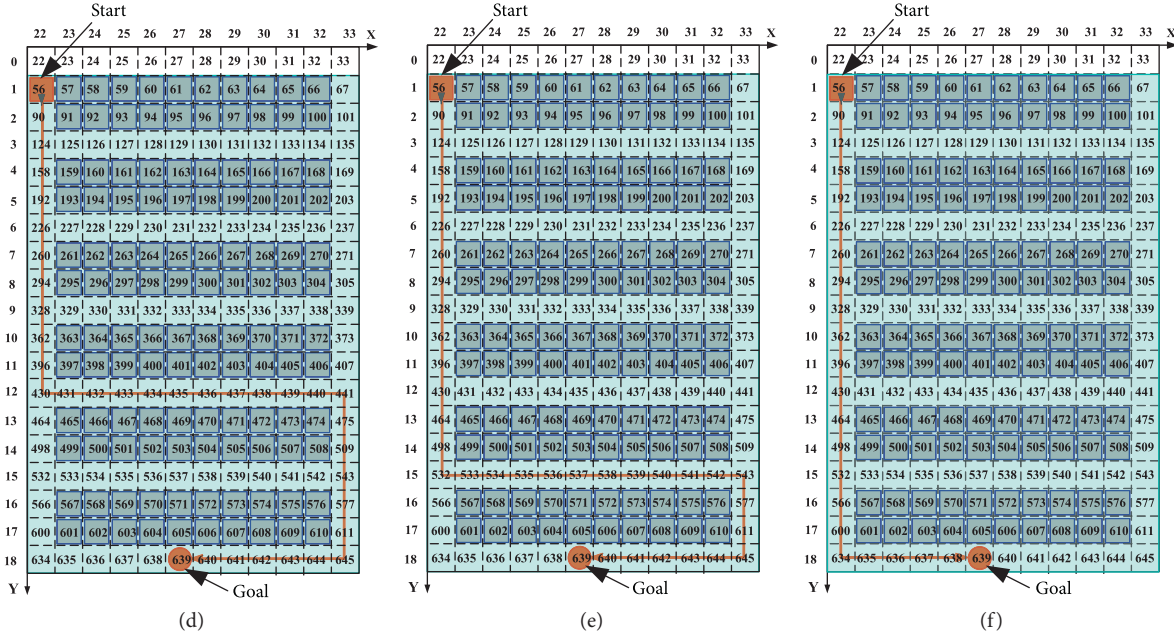


FIGURE 8: All feasible routes from AGV's start node to its goal node.

routes with the least AGV turning angle using (22) and obtain a route with the least turning nodes using equation (23), as shown in Figure 9. The comparison of route computing time between the algorithm proposed in this paper and the classical DP algorithm is shown in Figure 10, where the blue line represents the classical DP algorithm and the red the algorithm proposed in this paper.

As shown in Figure 10, the route computing time of the classical DP algorithm is **0.8756 s**, the time of the algorithm proposed in this paper is **0.1233 s**, and it reduces the route computing time (i.e., t_c) by nearly **85.92%**. Then, we compare the route computing time of the classical DP algorithm and our algorithm in SRNSs of different sizes (see Figure 11).

As shown in Figure 11, the algorithm proposed in this paper greatly reduces route computing time of AGVs and significantly improves route-planning algorithm's efficiency, and the larger the SRNSs, the more obvious the advantages of the proposed algorithm.

Case 2. Route planning for loaded AGVs and a no-load AGV in AS/RSs. Consider the start node of an AGV to be node 23 ($x=23, y=0$) and the goal node to be node 435 ($x=27, y=12$). If the AGV is loaded, it should avoid shelves located on nodes while traveling, and we need to build $obstacle_list = []$. Then, we compute route-planning regions using equations (8)–(13) (or equations (14)–(19)) and repeatedly employ equation (21) to obtain all the feasible routes for the AGV from the start node to the goal node. Finally, we obtain the route with the shortest travel time according to equations (22) and (23), as indicated by the red line in Figure 12. If the AGV is a no-load one, it does not need to avoid shelves while traveling. We omit

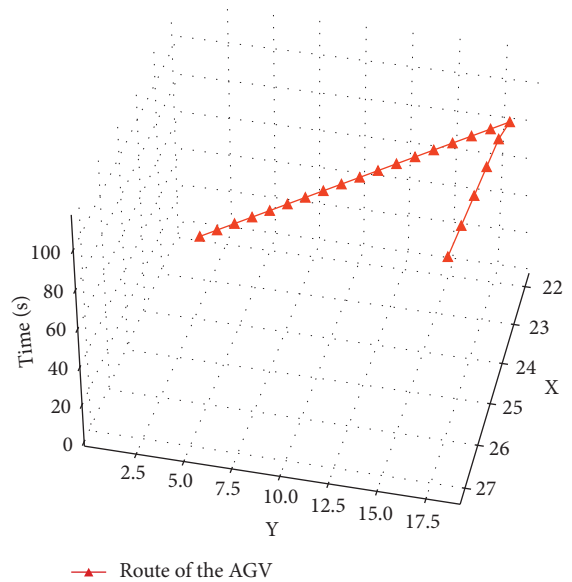


FIGURE 9: AGV's route that has the shortest distance and the least travel time.

the process of building $obstacle_list = []$ and directly use the region-segmentation DP algorithm to obtain the route with the shortest length and travel time, as the blue line shows in Figure 12.

As shown in Figure 12, although the start node and the goal node of loaded and no-load AGV are the same, the number of turns in routes, the length of the route, and the traveling time of the AGV are different, and the comparison results are as follows.

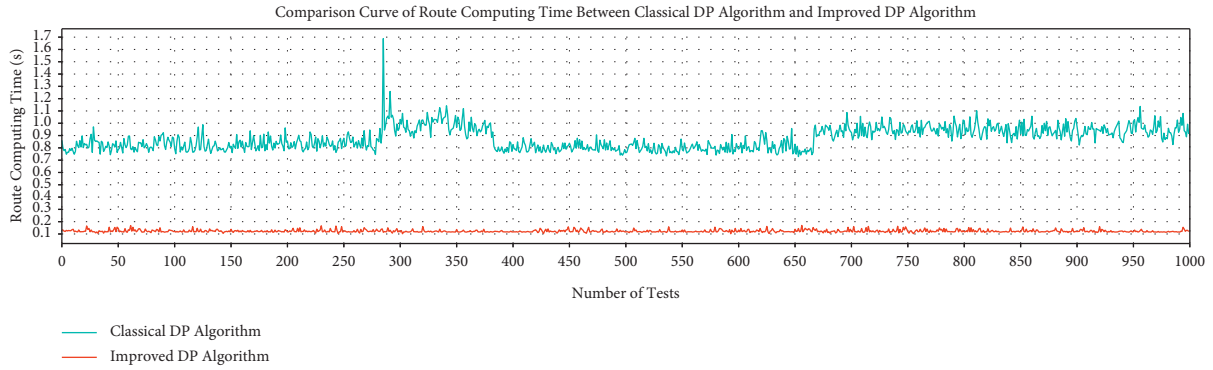


FIGURE 10: Comparison of route computing time between the classical DP algorithm and our algorithm.

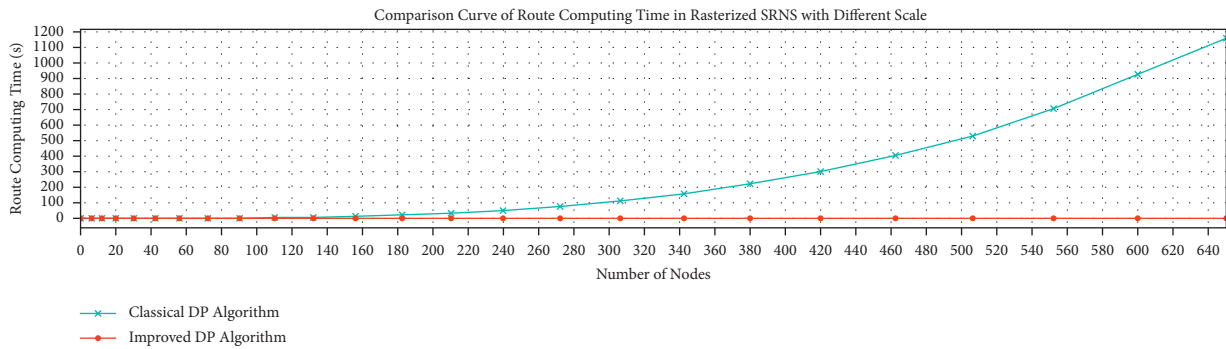


FIGURE 11: Comparison of route computing time between the classical DP algorithm and our algorithm.

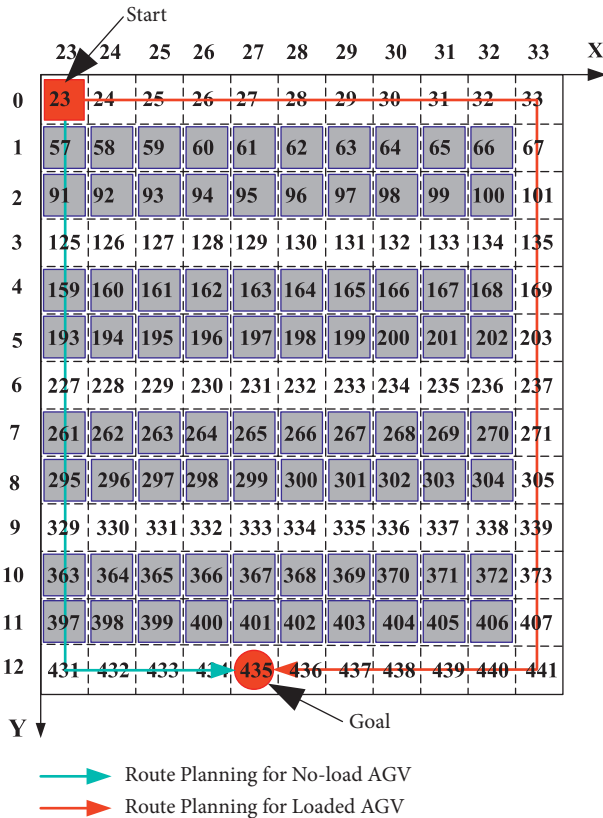


FIGURE 12: Route-planning results according to AGV's status (loaded or no-load).

TABLE 1: All feasible routes from AGV's start node to its goal node.

	Number of turns	Length of the route (m)	AGV's travel time (s)
Loaded AGV	2	70	150
No-load AGV	1	40	85

As shown in Table 1, (i) when the AGV is loaded, the number of turns is 2, the length of the route is 70m, and AGV's travel time is 150 s; (ii) when the AGV is a no-load one, the number of turns is 1, the length of the route is 40m, and AGV's travel time is 85s. In this scenario, the algorithm proposed in this paper improves the routes computing time (i.e., t_c) by nearly 43.33%.

5. Conclusions

This paper proposes an AGVs route-planning approach in large SRNSs based on a region-segmentation Dynamic Programming algorithm. First, we use the improved MDP to model the large SRNSs. Then, the large SRNSs are divided into several route-planning regions according to AGVs' start nodes and their goal nodes, the objective is to narrow the range of searching routes and reduce the number of samples, and this step significantly improved the route-planning speed. Since more travel time and energy cost can be caused by AGVs' turns, we compute the absolute value of angle deviation between each feasible route to AGVs' initial directions to choose the candidate routes, and then the route

with the least turns is selected as the optimal one. Compared with the conventional Dynamic Programming algorithm, the algorithm proposed in this paper greatly improves the efficiency of route planning in large SRNSs.

Data Availability

The data set used to support the simulation studies of this paper is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. Ning, G. Han, P. Duan, and J. Tan, "A global and dynamic route planning application for smart transportation," in *Proceedings of the First International Conference on Computational Intelligence Theory, Systems and Applications (CCITSA)*, December 2015.
- [2] K. Li, W. Ni, and E. Tovar, "On-board deep Q-network for UAV-assisted online power transfer and data collection," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12215–12226, 2019.
- [3] M. Ramasarayanan and M. Kirithikadevi, "End-to-End trust based transmission optimization in smartgrid network architecture," *International Journal of Computer Trends and Technology*, vol. 5, no. 3, 2013.
- [4] R. Katsuki, T. Tasaki, and T. Watanabe, "Graph search based local path planning with adaptive node sampling," in *Proceedings of the 2018 IEEE Intelligent Vehicles Symposium*, Changshu, China, June 2018.
- [5] D. Kularatne, S. Bhattacharya, and M. A. Hsieh, "Going with the flow: a graph based approach to optimal path planning in general flows," *Autonomous Robots*, vol. 42, no. 7, 2018.
- [6] E. Glorieux, P. Franciosa, and D. Ceglarek, "Coverage path planning with targeted viewpoint sampling for robotic free-form surface inspection," *Robotics and Computer-Integrated Manufacturing*, vol. 61, no. Feb., pp. 1–11, 2020.
- [7] X. Wang, Y. Yan, and X. Gu, "Spot welding robot path planning using intelligent algorithm," *Journal of Manufacturing Processes*, vol. 42, no. June, pp. 1–10, 2019.
- [8] N. Sugianti, A. Mardhiyah, and N. R. Fadilah, "Komparasi kinerja algoritma BFS, Dijkstra, greedy BFS, dan A* dalam melakukan pathfinding," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 5, no. 3, 2020.
- [9] K. Niemeyer and C. J. Sung, "DRGEP-based mechanism reduction strategies: graph search algorithms and skeletal primary reference fuel mechanisms," in *Proceedings of the 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, Orlando, Florida, July 2016.
- [10] S. Permana, K. Bintoro, B. Arifitama, and A. Syahputra, "Comparative analysis of pathfinding algorithms A*, Dijkstra, and BFS on maze runner game," *International Journal of Information System & Technology*, vol. 1, no. 2, 2018.
- [11] Q. Guo, Z. Zhang, and Y. Xu, "Path-planning of automated guided vehicle based on improved Dijkstra algorithm," in *Proceedings of the 29th Chinese Control and Decision Conf.*, pp. 7280–7285, Chongqing, China, May 2017.
- [12] R. Bellman, "Dynamic programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966.
- [13] R. I. Shreeve, "Introduction to dynamic programming, by leon cooper and mary W. Cooper. Pp 289. £15 hardback, £8.50 paperback. 1981. ISBN 0-08-0250645 (pergamon)," *The Mathematical Gazette*, vol. 66, no. 436, pp. 174–175, 1982.
- [14] D. Bertsekas, "Multiagent reinforcement learning: rollout and policy iteration," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 249–272, 2021.
- [15] A. G. Barto, "Reinforcement learning and dynamic programming," *Analysis, Design and Evaluation of Man-Machine Systems*, Elsevier Science, Amsterdam, Netherlands, pp. 407–412, 1995.
- [16] Y. D. Tian, "A simple analysis of AlphaGo," *Acta Automatica Sinica*, vol. 42, no. 5, pp. 671–675, 2016.
- [17] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, 1959.
- [18] A. R. Leach and A. P. Lemon, "Exploring the conformational space of protein side chains using dead-end elimination and the A* algorithm," *Proteins-structure Function & Bioinformatics*, vol. 33, no. 2, pp. 227–239, 2015.
- [19] T. Oral and F. Polat, "MOD* lite: an incremental path planning algorithm taking care of multiple objectives," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 245–257, 2015.
- [20] H. Akbaripour and E. Masehian, "Semi-lazy probabilistic roadmap: a parameter-tuned, resilient and robust path planning method for manipulator robots," *International Journal of Advanced Manufacturing Technology*, vol. 89, no. 5–8, pp. 1401–1430, 2016.
- [21] T. Xu, Y. Xu, D. Wang, S. Chen, W. Zhang, and L. Feng, "Path planning for autonomous articulated vehicle based on improved goal-directed rapid-exploring random tree," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–14, 2020.
- [22] M. Drust, T. Dietz, and A. Verl, "Dynamic and interactive path planning and collision avoidance for an industrial robot using artificial potential field based method," in *Proceedings of the 9th International Conference Mechatronics 2011*, Warsaw, Poland, September 2011.
- [23] Y. Zhang, S. Li, and H. Guo, "A type of biased consensus-based distributed neural network for path planning," *Nonlinear Dynamics*, vol. 89, pp. 1803–1815, 2017.
- [24] A. Tuncer and M. Yildirim, "Dynamic path planning of mobile robots with improved genetic algorithm," *Computers & Electrical Engineering*, vol. 38, no. 6, pp. 1564–1572, 2012.
- [25] G. Liu, H. Hou, and J. Liu, "Convergence analysis and improvement method of ant colony algorithm of path planning for mobile robot," in *Proceedings of the 2nd International Conference on Artificial Intelligence*, pp. 3839–3842, Management Science and Electronic Commerce (AIMSEC), Zhengzhou, China, August 2011.
- [26] L. He, P. Lou, X. Qiao, and R. Liu, "Conflict-free automated guided vehicles routing based on time window," *Computer Integrated Manufacturing Systems*, vol. 16, pp. 2630–2634, 2010.
- [27] Y. Yuan and L. Tang, "Novel time-space network flow formulation and approximate dynamic programming approach for the crane scheduling in a coil warehouse," *European Journal of Operational Research*, vol. 262, no. 2, pp. 424–437, 2017.
- [28] C. Novoa and S. Robert, "An approximate dynamic programming approach for the vehicle routing problem with stochastic demands," *European Journal of Operational Research*, vol. 196, no. 2, pp. 509–515, 2009.

- [29] M. Çimen and S. Mehmet, "Time-dependent green vehicle routing problem with stochastic vehicle speeds: an approximate dynamic programming algorithm," *Transportation Research Part D: Transport and Environment*, vol. 54, pp. 82–98, 2017.
- [30] H. Bahlawan, M. Morini, M. Pinelli, and P. R. Spina, "Dynamic programming based methodology for the optimization of the sizing and operation of hybrid energy plants," *Applied Thermal Engineering*, vol. 160, Article ID 113967, 2019.
- [31] T. Horiguchi, H. Takahashi, K. Hayashi, and C. Yamaguchi, "Dynamic programming for optimal packet routing control using two neural networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 339, no. 3-4, pp. 653–664, 2004.
- [32] M. W. Ulmer, J. C. Goodson, D. C. Mattfeld, and B. W. Thomas, "On modeling stochastic dynamic vehicle routing problems," *EURO Journal on Transportation and Logistics*, vol. 9, no. 2, Article ID 100008, 2020.
- [33] S. Desai and G. J. Lim, "Solution time reduction techniques of a stochastic dynamic programming approach for hazardous material route selection problem," *Computers & Industrial Engineering*, vol. 65, no. 4, pp. 634–645, 2013.
- [34] C. B. Browne, E. Powley, and D. Whitehouse, "A survey of Monte Carlo tree search methods," *IEEE Trans Comput Intell AI in games*, vol. 4, no. 1, pp. 1–43, 2012.
- [35] Z. Zhang, Q. Guo, J. Chen, and P. Yuan, "Collision-free route planning for multiple AGVs in an automated warehouse based on collision classification," *IEEE Access*, vol. 6, pp. 26022–26035, 2018.

Research Article

An Algorithm for Construction Project Cost Forecast Based on Particle Swarm Optimization-Guided BP Neural Network

Dan Ye 

School of Management Engineering, Jiangxi Electromechanical Vocational and Technical College, Nanchang 330038, China

Correspondence should be addressed to Dan Ye; yd0528@126.com

Received 1 July 2021; Revised 19 September 2021; Accepted 21 September 2021; Published 21 October 2021

Academic Editor: Wenbing Zhao

Copyright © 2021 Dan Ye. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Construction project cost prediction is an important function in construction-related fields; it can provide an important basis for project feasibility study and design scheme comparison and selection, and its accuracy will directly affect the investment decision of the project. The successful realization of construction cost prediction can bring great convenience to the control and management of construction cost. The purpose of this paper is to study a fast, accurate, convenient, deducible, and rational construction project cost prediction method, to provide a basis for the cost management of the whole life cycle of the project. Therefore, this paper uses particle swarm optimization algorithm to improve BP neural network and proposes a novel construction project cost prediction algorithm based on particle swarm optimization-guided BP neural network. Aiming at the defects of BP neural network updating weights and thresholds with the gradient descent method, this paper uses the advantages of particle swarm optimization in the field of parameter optimization to optimize BP neural network with PSO algorithm. The structure of BP neural network weights and the threshold of each neuron in the coding, through intelligent search for each particle, find the most suitable weights and thresholds, so that the BP neural network has faster convergence speed, better generalization ability, and higher prediction precision. Simulation results also show that the proposed algorithm is competitive enough.

1. Introduction

Construction project management [1–3] primarily consists of preliminary investment estimation [4, 5], plan design expansion design [6], construction drawing design [7], stage design budget, project budget in the bidding stage, project settlement, and project final accounts after completion, among other things. The investment estimation of construction costs [8–10] is the focus of construction project management. The profitability of a project is determined by the investment estimate of the construction cost [11]. The cost of construction and installation works, or the cost of construction works, plays a significant role in estimating the investment value of construction projects. As a result, estimating the cost of a construction project is crucial [12].

Predicting the cost of a construction project [13] is an important topic in the construction industry. It can serve as a solid foundation for project feasibility studies and design alternatives. Its precision has a direct impact on project

investment decisions. The successful implementation of construction project cost prediction can improve the control and management of construction project costs, making it a valuable research topic.

The cost of a construction project is forecasted using historical data from similar projects and mathematical models. In the past, traditional statistical analysis [14–16] and simple regression theory [17] were frequently used to predict construction project costs [18], using moving smoothing, linear regression, or the unit index method, for example. We discovered through extensive scientific research that the traditional construction engineering cost estimation method has several flaws, including low calculation accuracy and long calculation times. Since the influencing factors of construction project cost are many and complicated, and the data collected on the past project cost has high randomness and ambiguity. Also, when selecting indicators for estimation, most of the indicators have a certain degree of consistency. It refers to a geographical area

or the construction industry. The uniformity in the field, for example, did not take into account the management level, professional capabilities of the travel industry unit, as well as project quality, safety, and construction period, and was unable to adapt well to the market economic system. As a result, traditional construction project cost forecasting [19, 20] often fails to achieve satisfactory accuracy, and it often takes a long time. The construction project cost prediction loses its practical significance as a result of its low precision and time-consuming nature.

The realization of high-precision cost forecasting through mathematical modeling has piqued the interest of industry professionals and academics [21, 22], thanks to the rapid advancement of computer and neural network technologies [23–25]. Certain mathematical models and related historical engineering data are used to make the construction cost forecast. The BP neural network [26] is a relatively simple mathematical model in comparison to other learning models, but it has a lot of application value in project cost prediction. However, the traditional BP neural network prediction model has defects such as low calculation accuracy, poor stability, and insufficient generalization ability. Therefore, this paper intends to use the combination of particle swarm optimization algorithm and BP neural network to quickly predict the project cost.

The main contributions of this article are as follows:

- (1) This paper proposes a novel construction project cost prediction algorithm based on particle swarm optimization-guided BP neural network, which can predict the construction project cost more accurately and provide a basis for the cost management of the whole life cycle of the project.
- (2) This paper uses the advantages of particle swarm optimization in the field of parameter optimization to optimize the BP neural network through the PSO algorithm. That is, to code the weights and thresholds between the neurons in the BP neural network structure, and find the most suitable weights and thresholds through the intelligent search of each particle, so that the BP neural network has a faster convergence speed, stronger generalization ability, and higher prediction accuracy.

The rest of this article is organized as follows: Section 2 introduces the background of the research. Section 3 introduces the principle of the proposed algorithm in detail. Section 4 provides the experimental results. In Section 5, a conclusion based on this work is given.

2. Background

Construction project cost prediction is a very important aspect of work in the construction engineering industry, and it is very important in the management of construction projects. The prediction of construction project cost often occurs in the early stage of project construction. It is the basis of the feasibility study of construction projects and the important basis for the comparison and selection of design schemes, which will directly affect the investment decision of

the project. In view of the timely accuracy required for project investment decision-making, therefore, the accuracy and effectiveness of construction project cost forecasts are of vital importance.

In the past, the prediction of construction project cost was mainly achieved through the unit index method, that is, according to the characteristics, structure, and scale of the project, the corresponding forecasting index was applied, calculated, and summarized. The whole process is relatively complicated and time consuming; at the same time, this set of forecasting schemes also has the problem that the forecasting accuracy is difficult to guarantee. The unit index method is used to predict the construction project cost. The index system is unified by the local or industry, each project's construction management level and construction site conditions are not uniform, and they have strong individualism. Therefore, the use of the unit index method fails to fully consider the individuality of each single project, which leads to the insufficient accuracy of the method for predicting the construction project cost.

In recent years, many domestic and foreign experts and scholars have proposed a new method of construction project cost prediction in order to achieve accurate and rapid construction cost forecasting, that is, forecasting based on traditional statistical analysis methods. Most of these researchers have adopted methods such as probability theory and linear regression to quickly predict the cost of construction projects. These forecasting technologies can often realize the forecast of construction project cost under certain conditions; but, they often have problems such as low forecasting accuracy and time-consuming forecasting, and their generalization ability is poor, and their application prospects are not good.

Many researchers are focusing their efforts on construction cost prediction, and it is difficult to find a more suitable mathematical model to achieve satisfactory construction cost prediction accuracy and speed. The previous unit index method had an index system that was uniformly formulated by localities or industries and had poor marketability. It is often difficult to consider the unique construction management circumstances of each project, and it has no promotional value. Traditional statistical analysis methods, such as linear regression analysis, are difficult to use and time consuming. In recent years, with the advancement of science and technology, the development of computer science and artificial intelligence theory, some intelligent mathematical theoretical models have been gradually applied to the application of these theories in the construction cost forecasting, and certain effects have been achieved. For example, with the development of the construction project cost prediction model based on artificial neural network prediction with higher prediction accuracy, the prediction speed is also faster. However, these research results are often due to the unreasonable construction of the project cost index system or the lack of clear selection of case projects, resulting in poor universality of the research results. There are still many problems to be solved in the prediction of construction project cost.

3. Methodology

3.1. Project Costs. The estimated or actual expenditure of the construction project during the construction period is referred to as project cost. In today's market, project cost can take on two different meanings depending on the supply and demand objects. The total fixed asset investment cost of the expected or actual expenditure of a project is analyzed from the perspective of the investor (owner). Investors must complete a series of activities, including investment decision-making, survey and design, bidding, construction, and completion and acceptance, in order to obtain the expected benefits of an investment project. The project cost refers to the total cost of the above activities. From this perspective, the project cost is the total investment in fixed assets of the construction project, as shown in Figure 1.

From the standpoint of market exchanges, the contract price of construction projects formed by the market is referred to as project cost. It is a common and important type of project cost. Using the construction project as the transaction object, the market forms the price agreed upon by the main body of demand (investors) and the main body of supply (contractors)—construction and installation project costs, through the contract transaction (most contract parties bid, and the contractor bids), on the basis of multiple estimates.

From the above content, we can see that the meaning of the project cost changes with the change of the construction object. In fact, the two meanings grasp the essence of the project cost from different perspectives. For the investor (owner), the project cost not only involves the construction of a project to pay all the costs, in addition to the contract price, but also includes the construction of early investment management fees, consulting fees, research and test fees, environmental impact assessment fees, reserve fees, and interest generated in the construction period and other costs. But for the contractor, the cost of the construction project refers to the cost of the construction and installation project signed with the investor (owner). For the whole construction project, the cost of construction and installation can account for 65–70% of the total cost, and other costs are calculated based on the cost of construction and installation. Therefore, this paper takes the second meaning of construction installation project cost as the research object.

It is the construction and installation project investment in the construction project investment, as well as a component of the project cost, from the perspective of investment. It is a price determined by the market and agreed upon by the investor and the construction party in the context of market transactions. The cost of construction and installation projects refers to the costs of the project's construction, supporting projects, and production equipment installation. The structure is depicted in Figure 2.

3.2. Particle Swarm Optimization Algorithm

3.2.1. Mathematical Description of PSO. The goal of particle swarm optimization (PSO) is to set the initial position and speed of a swarm of random particles, then find the best solution using a constant iterative search under certain conditions. It is a clever algorithm that is based on group behavior. Bird behavior appears to be under some control, according to research into their flight characteristics. There appears to be some relationship between individuals and between individuals and groups under this control, and the birds rely on this connection for food.

Suppose that in a D -dimensional target search space, there are n particles that represent the possible solutions of the problem, that is to say, the position of the particle is the possible solution of the research problem in the D -dimensional space. The population is $X = \{x_1, x_2, \dots, x_d\}$, where the position and velocity of the particle i are $X_i = \{x_{i1}, x_{i2}, \dots, x_{id}\}$, $i = 1, 2, \dots, n$, $V_i = \{v_{i1}, v_{i2}, \dots, v_{id}\}$, $i = 1, 2, \dots, n$ and the particles will be updated in the iterative process according to the optimal solution. One is the best position P_{best} of the individual searched by the particle, that is, $P_i = \{p_{i1}, p_{i2}, \dots, p_{id}\}$, $i = 1, 2, \dots, n$, and the other one is the best position G_{best} determined by the population search, that is, $P_g = \{p_{g1}, p_{g2}, \dots, p_{gd}\}$, $g = 1, 2, \dots, n$. When the i -th particle finds the abovementioned best extreme point, the velocity of the particle and the position of the next iteration are updated according to the following equation:

$$\begin{aligned} v_{id}^{k+1} &= v_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k), \\ X_{id}^{k+1} &= X_{id}^k + v_{id}^{k+1}, \end{aligned} \quad (1)$$

where k is the number of iterations and c_1 and c_2 are acceleration constants, namely, learning factors. c_1 mainly adjusts the step size of the particle flying to its own best position, c_2 mainly adjusts the step size of the particle flying to the global best position, r_1 and r_2 are random numbers evenly distributed between 0 and 1.

In order to improve the search ability of particles, a particle swarm algorithm with inertia weight is proposed, and the inertia weight factor ω is introduced to the initial velocity of each iteration of the particles:

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k). \quad (2)$$

The introduction of inertial weight can affect the optimization ability of particles. Specifically, a larger ω can improve the global search ability of the algorithm, and a smaller ω can improve the local search ability of the algorithm. According to the different forms of inertial weight ω , a variety of particle swarm algorithms have been produced. Common ones include adaptive weights, linearly decreasing weights, and random weight algorithms. This paper uses linearly decreasing inertia weights, and the weight change formula is as follows:

$$\omega = \omega_{\max} - \frac{t * (\omega_{\max} - \omega_{\min})}{t_{\max}}. \quad (3)$$

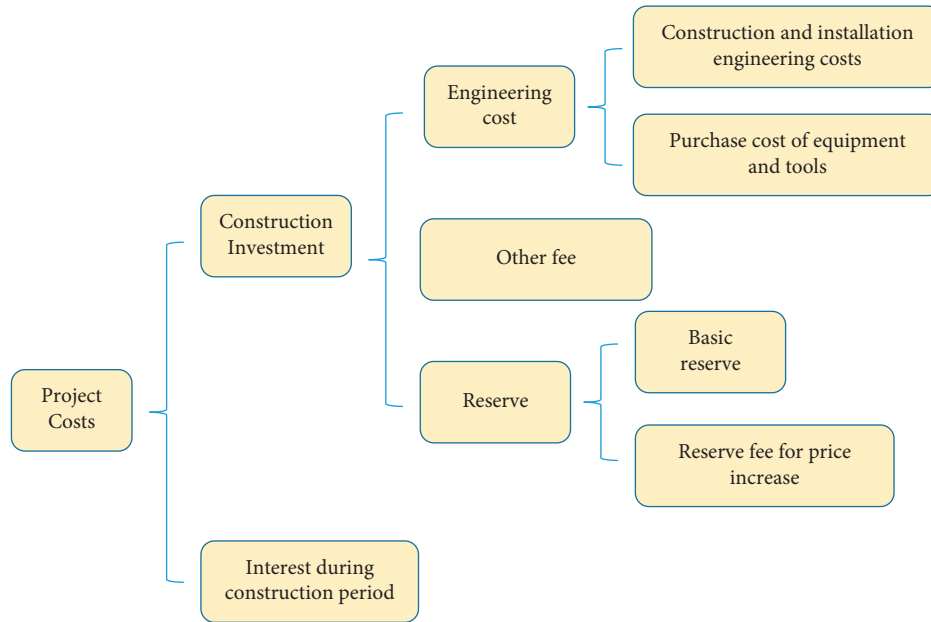


FIGURE 1: Composition of the project cost.

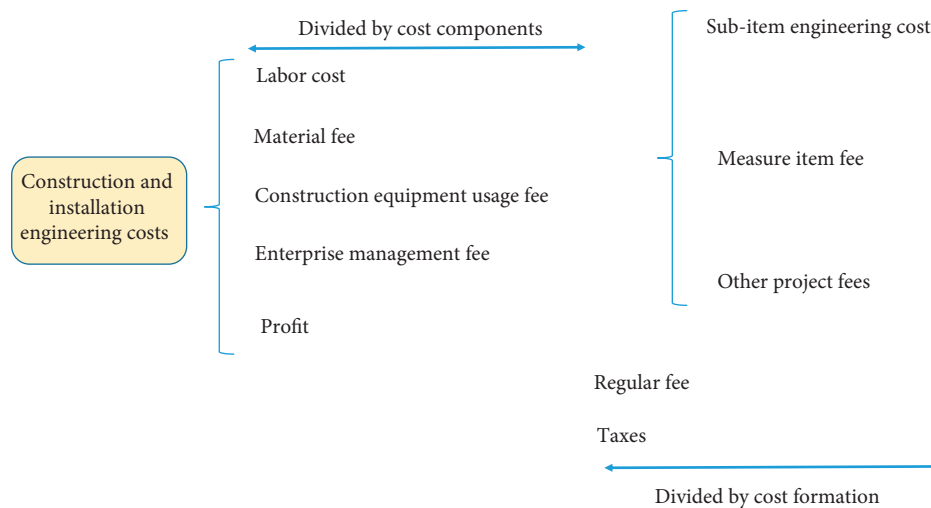


FIGURE 2: Composition of construction and installation engineering costs.

3.2.2. *Optimization Process of PSO.* The PSO algorithm flow is shown in Algorithm 1.

3.3. *Project Cost Prediction Model.* BP neural network algorithm is a single point search method based on gradient descent of error function, which has no global search ability. Therefore, in the process of learning and training of BP neural network, it is inevitable that there will be poor robustness, slow convergence speed, poor generalization ability, and other shortcomings. However, the PSO algorithm has the advantages of simple structure, large search range, strong robustness, fast convergence, and so on, and can solve most global optimal solutions. Therefore, this study combines the advantages of the two and establishes a

project cost prediction model based on PSO optimization and BP neural network parameters.

In the modeling process of the BP neural network, two key numbers need to be set: weight ω and threshold θ , which are a group of random values and easily fall into local minimum values. The training of weights and thresholds is actually a complex problem to find the optimal parameters. The gradient descent method is used to update the connection weights and thresholds. Even if there is a slight change in weights and thresholds, the neural network will get completely different operating results. The determination and optimization of these two factors determine the generalization ability and stability of the model to a great extent. To optimize the BP network through the PSO algorithm, it is necessary to encode the weights and thresholds among each

Steps of specific operation

Step 1: set parameters such as ω , N , $c1$ and $c2$, termination conditions in the algorithm.

Step 2: initialize the population, including random positions and speeds.

Step 3: evaluate the fitness value of the particle fitness.

Step 4: the fitness value of each particle is compared with the best position (individual extreme value) it has passed. If the current fitness value is better, its position is taken as the current best position P_{best} .

Step 5: similarly, the fitness value of each particle is compared with the global best position it has passed, and if the current fitness value is better, its location is taken as the current global best position G_{best} .

Step 6: update the speed and position of particles.

Step 7: determine whether the termination conditions are met, if not, return to the third step to continue the iterative update.

Otherwise, the P_{best} corresponding to the current fitness value is output as the global optimal solution, and the search is stopped.

ALGORITHM 1: Flow of particle swarm algorithm.

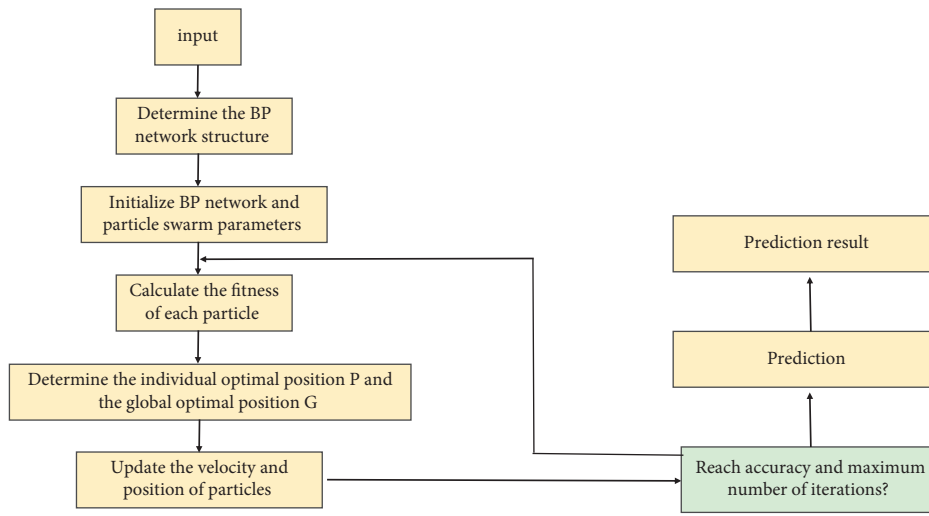


FIGURE 3: Schematic diagram of the model of construction project cost prediction algorithm based on particle swarm optimization-guided BP neural network.

neuron in the structure of the BP network, and search each particle intelligently to find the most appropriate weights and thresholds, so that the BP neural network has faster convergence speed and better generalization ability.

The individual in the PSO algorithm corresponds to the initial weight parameters of each layer of the BP network, and the weight coefficient of each layer is adjusted by the particle swarm algorithm. When the termination condition is met, the search is stopped. The structure of the BP neural network proposed in this paper is $i - b - j$. There are $i * b$ weights connecting the input layer and the hidden layer, $b * j$ weights connecting the hidden layer and the output layer, and the number of hidden layer thresholds is b . There are j thresholds in the output layer, so the dimension of the particle is $d = i * b + b * j + b + j$. The fitness function F of PSO is expressed in the following equation:

$$F = \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^N (y_{ij} - x_{ij})^2, \quad (4)$$

where M is the total number of input learning samples, N is the number of output nodes, y_{ij} is the actual output value of

TABLE 1: Hyperparameter setting.

.trainParam.goal = 0.1
.trainParam.epochs = 300
.trainParam.show = 20
.trainParam.mc = 0.95
.trainParam.lr = 0.05
.trainParam.min_grad = $1e - 6$
.trainParam.min_fail = 5

the corresponding parameter, and x_{ij} is the expected output value of the corresponding parameter. Therefore, the model of the construction project cost prediction algorithm based on particle swarm optimization-guided BP neural network is shown in Figure 3.

4. Experiments and Results

4.1. Experimental Setup. The Matlab 2018b platform is used to simulate and analyze the construction project cost prediction model developed in this study. It is created using a matrix system. Its instruction expression is similar to

TABLE 2: Examples of data sets.

ID	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	Y
1	52364.12	6321.23	31	2	2.8	95.3	4	104	0	4	3	1	1	3	1	4	4	3125.21
2	71256.23	4215.74	31	1	2.9	100	4	105	1	4	3	2	1	3	2	4	4	2369.23
...
227	25456.39	1423.05	33	3	2.9	99.3	4	108	2	3	3	1	1	3	1	3	4	3253.24

TABLE 3: Comparison of prediction results of different algorithms.

Test sample	True value	Predictive value		
		BP	ARIMA	Ours
214	2263.21	2305.32	2301.25	2271.23
219	2314.20	2412.23	2501.25	2319.22
220	2514.63	2519.65	2501.36	2515.23
222	2412.22	2432.33	2438.52	2410.25
225	2272.58	2289.25	2279.01	2274.36

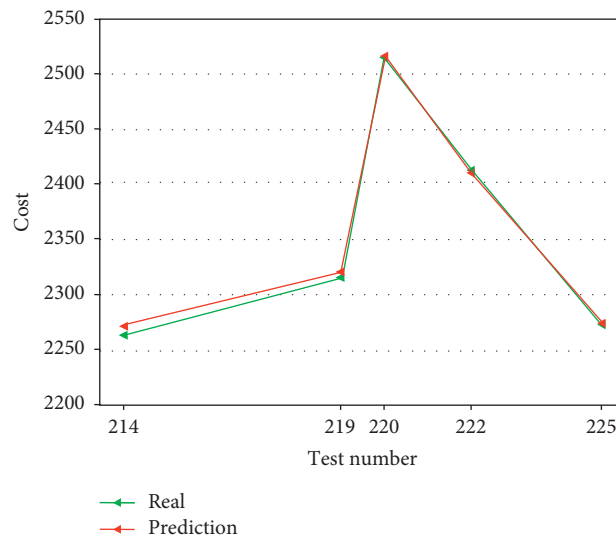


FIGURE 4: Forecast curve.

engineering mathematics language, reducing the time-consuming challenges connected with programming in language and other languages. It greatly reduces the number of programming sentences and improves the efficiency of computing. It is a type of powerful mathematical software that is well-suited for mathematical modeling and the processing of extremely complex mathematical operations. A large number of toolboxes are preinstalled, and users can name them as needed. In addition, the operating system used is Windows 10, and the parameter settings of the improved BP neural network are shown in Table 1.

4.2. Dataset. The data selected in this paper are from the final accounting data of a real estate enterprise's existing project in Jiangsu. We obtained the final account data of 240 completed high-rise residential projects constructed by the real estate company in Jiangsu in the past three years. After eliminating unnecessary and redundant information, 227

sets of valid training samples were obtained. In Section 3.1, the high-rise residential project cost prediction index includes numerical quantitative index and character qualitative index. For the selected sample data, quantitative indicators such as floor area and number of floors can be directly input according to the actual engineering data, but qualitative indicators such as basic types and interior wall decoration cannot be directly input. It makes it simple for the prediction model to learn and train on the sample data. The unilateral cost Y is taken as the output set, and the remaining indicators $X_1 X_{17}$ are taken as the input set. The data of qualitative indicators of character type after quantitative processing are shown in Table 2.

4.3. Evaluation Index. In this paper, the relative error and the relative error of the average absolute value are used to reflect the influence of different algorithms on the prediction effect of the model. The calculation equation is as follows:

TABLE 4: Comparison of prediction results of different algorithms.

Test sample	True value	Predictive value		
		BP-NO-PSO	PSO	BP + PSO (ours)
214	2263.21	2411.31	2357.88	2271.23
219	2314.20	2198.66	2411.44	2319.22
220	2514.63	2318.13	2688.11	2515.23
222	2412.22	2333.67	2422.66	2410.25
225	2272.58	2411.99	2316.43	2274.36

$$\delta = \frac{T_i - F_i}{T_i}, \quad (5)$$

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^m |\delta| \times 100\%,$$

where T_i and F_i represent the actual value and predicted value of the i -th sample, respectively, and m represents the number of test samples.

4.4. Forecast Result Analysis. According to the input set of the prediction model obtained above, 20% of the 227 sets of data were first selected as test data, and then the remaining 80% of the data were used as training data to train the model. BP neural network and PSO-BP neural network are used to simulate and predict the construction cost, respectively. The prediction comparison results are shown in Table 3.

The prediction model that optimizes the BP neural network parameters through the PSO algorithm is significantly better than the single BP neural network model in terms of predictive stability, as shown in Table 3 and Figure 4, and the model is more stable. Simultaneously, the relative errors of the three models can be controlled within 10%, and the prediction effects of the three models are excellent, as shown in Table 3. Forecast accuracy can be achieved during the investment decision-making stage. The optimized PSO-BP model, on the other hand, is significantly less accurate than the BP neural network, and the BP neural network's error is lower than the ARIMA model, as determined by the average absolute relative error of the test samples. Thus, the high-rise residential project cost prediction model based on PSO-BP neural network performs better in terms of error control and prediction accuracy.

The results show that the BP neural network model with optimized parameters has a good application effect in cost prediction. For the construction project cost prediction, the prediction model based on PSO optimization BP neural network parameters has a good guiding significance, and it is very suitable for the preliminary construction cost prediction.

4.5. Ablation Experiments. In order to further verify the effectiveness and superiority of the algorithm in this section, this section conducts the particle swarm optimization algorithm to improve the ablation experiment of the BP neural network. Let "PSO" represent the particle swarm algorithm, the results of the ablation experiment are shown in Table 4.

It can be clearly seen from Table 4 that the average absolute percentage error of using only the BP network is the

largest. Secondly, the MAPE of PSO is lower than that of the BP network, but no matter what, BP + PSP has achieved the best prediction performance. This again proves the superiority of the algorithm in this paper.

5. Conclusion

In this paper, a fast, accurate, convenient, and reasonable construction project cost forecasting method to provide a basis for the cost management of the whole life cycle of the project has been studied. Therefore, this paper improves the BP neural network through the particle swarm optimization algorithm, and proposes a novel construction project cost prediction algorithm based on the particle swarm optimization-guided BP neural network. Aiming at the defects of the BP neural network using gradient descent method to update weights and thresholds, this paper uses the advantages of the particle swarm algorithm in the field of parameter optimization to optimize the BP neural network through PSO algorithm. That is, to code the weights and thresholds between the neurons in the BP neural network structure, and find the most suitable weights and thresholds through the intelligent search of each particle, so that the BP neural network has a faster convergence speed. Stronger generalization ability, higher prediction accuracy, and simulation experiments also show that the discussed algorithm has competitive performance.

In addition, considering that the current work is time series data, in the next work, an attempt will be made to use the LSTM network to continue the research.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares no conflicts of interest.





References

- [1] S. Demirkesen and B. Ozorhon, "Impact of integration management on construction project management performance," *International Journal of Project Management*, vol. 35, no. 8, pp. 1639–1654, 2017.
- [2] S. Banihashemi, M. R. Hosseini, H. Golizadeh, and S. Sankaran, "Critical success factors (CSFs) for integration of sustainability into construction project management practices in developing countries," *International Journal of Project Management*, vol. 35, no. 6, pp. 1103–1119, 2017.

- [3] S. Kim, S. Chang, and D. C. Lacouture, "Dynamic modeling for analyzing impacts of skilled labor shortage on construction project management," *Journal of Management in Engineering*, vol. 36, no. 1, Article ID 04019035, 2020.
- [4] P. Mesároš and T. Mandičák, "Exploitation and benefits of BIM in construction project management," in *Proceedings of the IOP Conference Series: Materials Science and Engineering*, vol. 245, no. 6, October 2017, Article ID 062056.
- [5] J. Symons, E. Howard, K. Sweeny, M. Kumnick, and P. Sheehan, "Reduced road traffic injuries for young people: a preliminary investment analysis," *Journal of Adolescent Health*, vol. 65, no. 1, pp. S34–S43, 2019.
- [6] T. Gopalakrishnan and R. Saravanan, "Cast off expansion plan by rapid improvement through Optimization tool design, Tool Parameters and using Six Sigma's ECRS Technique," in *Proceedings of the IOP Conference Series: Materials Science and Engineering*, vol. 183, no. No. 1, October 2017, Article ID 012016.
- [7] V. Senthilkumar and K. Varghese, "Structured methodology to formulate drawing dependency structure matrix for construction design," *Architectural Engineering and Design Management*, vol. 5, no. 4, pp. 225–248, 2009.
- [8] H. Adeli and M. Wu, "Regularization neural network for construction cost estimation," *Journal of Construction Engineering and Management*, vol. 124, no. 1, pp. 18–24, 1998.
- [9] M. Juszczak, "The challenges of nonparametric cost estimation of construction works with the use of artificial intelligence tools," *Procedia engineering*, vol. 196, pp. 415–422, 2017.
- [10] J. Ahn, S. H. Ji, and S. J. Ahn, "Performance evaluation of normalization-based CBR models for improving construction cost estimation," *Automation in Construction*, vol. 119, Article ID 103329, 2020.
- [11] C. A. Peleskei, V. Dorca, R. A. Munteanu, and R. Munteanu, "Risk consideration and cost estimation in construction projects using Monte Carlo simulation," *Management*, vol. 10, no. 2, 2015.
- [12] A. Firouzi, W. Yang, and C.-Q. Li, "Prediction of total cost of construction project with dependent cost items," *Journal of Construction Engineering and Management*, vol. 142, no. 12, Article ID 04016072, 2016.
- [13] C. Koo, T. Hong, and C. Hyun, "The development of a construction cost prediction model with improved prediction capacity using the advanced CBR approach," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8597–8606, 2011.
- [14] D. S. Tejale, S. D. Khandekar, and J. R. Patil, "Analysis of construction project cost overrun by statistical method," *International Journal*, vol. 3, no. 5, pp. 349–355, 2015.
- [15] A. A. Abu Hammad, S. M. A. Ali, G. J. Sweis, and R. J. Sweis, "Statistical analysis on the cost and duration of public building projects," *Journal of Management in Engineering*, vol. 26, no. 2, pp. 105–112, 2010.
- [16] C.-h. Wu, T.-y. Hsieh, and W.-l. Cheng, "Statistical analysis of causes for design change in highway construction on Taiwan," *International Journal of Project Management*, vol. 23, no. 7, pp. 554–563, 2005.
- [17] T. P. Williams, "Predicting final cost for competitively bid construction projects using regression models," *International Journal of Project Management*, vol. 21, no. 8, pp. 593–599, 2003.
- [18] D. Baccarini, "Accuracy in estimating project cost construction contingency—a statistical analysis," in *Proceedings of the Cobra 2004: RICS International Construction Conference, Responding to Change*, Leeds, United Kingdom, September 2004.
- [19] H. Leon, H. Osman, M. Georgy, and M. Elsaid, "System dynamics approach for forecasting performance of construction projects," *Journal of Management in Engineering*, vol. 34, no. 1, Article ID 04017049, 2018.
- [20] S. Bayram and S. A. Jibouri, "Efficacy of estimation methods in forecasting building projects' costs," *Journal of Construction Engineering and Management*, vol. 142, no. 11, Article ID 05016012, 2016.
- [21] M. Mir, H. M. D. Kabir, F. Nasirzadeh, and A. Khosravi, "Neural network-based interval forecasting of construction material prices," *Journal of Building Engineering*, vol. 39, Article ID 102288, 2021.
- [22] K. Bala, S. A. Bustani, and B. S. Waziri, "A computer-based cost prediction model for institutional building projects in Nigeria: an artificial neural network approach," *Journal of Engineering, Design and Technology*, vol. 12, no. 4, 2014.
- [23] S. H. Iranmanesh and M. Zarezadeh, "Application of artificial neural network to forecast actual cost of a project to improve earned value management system," in *Proceedings of the World congress on Science, Engineering and Technology*, pp. 240–243, São Paulo, Brazil, March 2008.
- [24] M. Zhao, C. H. Chang, W. Xie, Z. Xie, and J. Hu, "Cloud shape classification system based on multi-channel cnn and improved fdm," *IEEE Access*, vol. 8, Article ID 44111, 2020.
- [25] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local-global multiple cfrvtfkr," *Sensors*, vol. 21, no. 4, 2021.
- [26] L. Huang, G. Xie, W. Zhao, Y. Gu, and Y. Huang, "Regional logistics demand forecasting: a bp neural network approach," *Complex & Intelligent Systems*, pp. 1–16, 2021.

Research Article

Load Forecasting Method Based on Improved Deep Learning in Cloud Computing Environment

Kai Zhang ¹, Wei Guo ², Jian Feng ¹ and Mei Liu ³

¹State Grid Hebei Electric Power Co., Ltd., Shijiazhuang 050021, China

²Market Service Center, State Grid Hebei Electric Power Co., Ltd., Shijiazhuang 050021, China

³Beijing Tsingsoft Technology Co., Ltd., Beijing 100085, China

Correspondence should be addressed to Wei Guo; yxzx_guow@he.sgcc.com.cn

Received 4 August 2021; Revised 8 September 2021; Accepted 11 September 2021; Published 12 October 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Kai Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For the problems of low accuracy and low efficiency of most load forecasting methods, a load forecasting method based on improved deep learning in cloud computing environment is proposed. Firstly, the preprocessed data set is divided into several data partitions with relatively balanced data volume through spatial grid, so as to better detect abnormal data. Then, the density peak clustering algorithm based on spark is used to detect abnormal data in each partition, and the local clusters and abnormal points are merged. The parallel processing of data is realized by using spark cluster computing platform. Finally, the deep belief network is used for load classification, and the classification results are input into the empirical mode decomposition-gating recurrent unit network model, and the load prediction results are obtained through learning. Based on the load data of a power grid, the experimental results demonstrate that the mean prediction error of the proposed method is basically controlled within 3% in the short term and 0.023 MW, 19.75%, and 2.76% in the long term, which are better than other comparison methods, and the parallel performance is good, which has a certain feasibility.

1. Introduction

With the vigorous development of social and economic level, the power consumption of industry, commerce, and residents in the power grid shows the characteristics of rapid growth. Due to the different production processes and peak and valley periods of power consumption among different industries, there are certain differences in load characteristics among different users [1, 2]. At the same time, the load characteristics of power users also change with seasonal changes, weather changes, characteristic days, power consumption areas, and other factors. Some loads fluctuate violently in a short time and there is a great peak valley difference [3]. Accurately analyze the load characteristics and make corresponding prediction, which can not only improve the power supply economic benefits of the power grid but also help the power grid provide strong decision support in optimizing energy structure and rational allocation of resources [4, 5]. Therefore, how to

comprehensively analyze the power load characteristics to achieve accurate load forecasting has become an important and difficult research.

On the other hand, due to the extensive use of sensors and smart meters in the power grid, it can quickly obtain massive power consumption data with high dimension and fine granularity, thus forming user power big data. The load big data contains rich user characteristic information, which can be deeply mined to give full play to its maximum value [6, 7]. At present, the commonly used load data analysis and application methods can include two types. The first type of method is time series analysis method, such as multiple linear regression, autoregressive moving average model, and other methods. This method requires that the time series used for calculation be relatively stable. It is generally suitable for long-term load forecasting with stable growth, and it is difficult to be applied to short-term load with frequent fluctuations [8, 9]. The second type is artificial intelligence technology represented by neural network and

its improved combination method. This method has good overall prediction performance but does not consider the characteristics of different time periods and different load types, and its efficiency needs to be improved when dealing with huge data volume and complex data structure [10].

In view of the unsatisfactory performance of existing methods in power grid load forecasting, a power grid load forecasting method based on improved deep learning in cloud computing environment is proposed. Its innovations are summarized as follows:

- (1) Due to the high time complexity of the fast density peak algorithm, the proposed method divides the data set to be clustered into multiple data partitions with relatively balanced data volume through spatial grid and designs a parallel algorithm by using spark parallel programming model. The power load is detected in parallel in the data partition corresponding to each computing node, and the detected abnormal point set is combined, so as to reduce the computational complexity and ensure the accuracy of abnormal data detection.
- (2) For full use of the massive data of power grid load and considering the periodicity and regularity of load itself, the proposed method finds the commonness between load data through fuzzy c-means clustering algorithm and inputs it into deep belief network (DBN) to classify the daily load to be measured.
- (3) For the high nonlinearity and instability of power load series caused by the superposition of various influencing factors, the proposed method combines empirical mode decomposition (EMD) and gated recurrent unit (GRU) to predict load. It avoids the random errors caused by modeling and forecasting the decomposed subsequences, respectively, so as to predict the load more accurately and quickly.

2. Related Research

Load forecasting is a process to predict the future load change and explore the dynamic and internal law of load data by analyzing the historical load series [11]. In recent years, academia has carried out a lot of research work on load forecasting methods. Most of the existing research results include three categories: traditional methods, artificial intelligence methods, and combined prediction methods [12].

Among them, traditional load forecasting method includes time series method and multiple linear regression method [13], which can show good performance in processing general load data, but when analyzing high-dimensional and complex power load data, the prediction accuracy is lacking, and the analysis efficiency is not high. The artificial intelligence and combination method play a great role in improving the performance of power grid load forecasting. Artificial intelligence methods include gray theory and support vector machine. For example, [14] constructs a gray correlation analysis model based on

interval gray effective information transformation, optimizes the resolution of traditional gray, and puts forward a multi-variable gray model to predict interval gray series, so as to effectively deal with the data prediction problem in big data. Reference [15] proposed a load forecasting method based on deep learning method for heat load demand, which has higher accuracy than the linear model of automatic feature selection, but the amount of calculation is relatively large. Reference [16] proposed a sequence to sequence recurrent neural network with attention to power load forecasting to capture the time dependence in load data. When dealing with huge and complex load data, a single artificial intelligence method can achieve good analysis performance in one aspect, but the overall load forecasting performance is not ideal.

The development of artificial intelligence technology promotes the continuous optimization of load forecasting. The combined forecasting method includes model combination in forecasting mechanism and weighted combination of forecasting results [17]. For example, [18] proposed two data preprocessing methods based on wavelet transform to extract data features, combined with least squares support vector machine learning engine and improved virus colony search algorithm to achieve accurate load prediction, but the prediction time range is limited. Reference [19] proposed a load forecasting method based on gated recurrent unit (GRU) combined with deep learning idea. Based on deep learning, different types of load influencing factors are processed, and the gated cyclic unit network is introduced to process the historical load series with time series characteristics, so as to finally complete the load forecasting. However, this method depends on data timing and needs more data preprocessing. Reference [20] proposed a short-term load forecasting combination method combining fuzzy time series and convolutional neural networks (CNN). By using the images created by the sequence values of multivariate time series and combined with CNN model, the relevant important parameters are automatically determined and extracted to accurately complete the load forecasting. Reference [21] proposed long-term and short-term memory, multilayer perceptron, and CNN to learn the relationship in the time series, but it is highly dependent on the load data in terms of capture time. Reference [22] realizes fast and accurate short-term load forecasting based on stacking factor condition limited Boltzmann machine and condition limited Boltzmann machine, but the model is complex and the processing efficiency is poor. The role of cloud computing environment on big data processing efficiency has not been deeply considered.

To sum up, the existing methods are difficult to give consideration to both prediction efficiency and accuracy in the process of load forecasting. Therefore, a power grid load forecasting method based on improved deep learning algorithm in cloud computing environment is proposed. In spark cluster computing platform, fuzzy C-means clustering is used to mine the relationship between data, DBN is used to realize load classification, and the classified load information is input into EMD-GRU combination model to complete load forecasting.

3. Proposed Method

3.1. Overall Framework. Spark is a fast and general cluster computing platform. Its main feature is that it can perform operations in memory and has high processing efficiency. Under the spark framework, combined with the cluster analysis method in data mining technology, a load forecasting method based on improved deep learning is proposed, and its process is shown in Figure 1.

Firstly, the preprocessed load data is partitioned through spark to better detect abnormal data. Then, in each partition, the density peak clustering algorithm is used to detect abnormal data, eliminate bad data, and reduce the impact of data on load forecasting. At the same time, combined with the main influencing factors and typical load types, DBN is used to classify different loads and determine the load category for forecast days. Finally, the load type and original load for forecast days are input into the EMD-GRU model for learning and training, so as to quickly match the corresponding model to realize load forecasting.

3.2. Load Data Preprocessing. In the power grid, the load data is measured by various sensors, and individual data is often lost or distorted due to collection, transmission, storage, and other factors. Before starting the prediction, the data needs to be preprocessed [23].

For the processing of missing data, the average method is generally used to fill in the missing value. The calculation is as follows:

$$L_t = \frac{L_{t-1} + L_{t+1}}{2}, \quad (1)$$

where L_t is the filling value at time t ; L_{t-1} and L_{t+1} represent the load values of the previous time and the next time, respectively. When the data at the beginning and end of the sequence is missing, the trend extrapolation method can be used for completion; and it is necessary to distinguish the authenticity of data values to improve the accuracy of prediction, that is, to identify and correct abnormal data. Abnormal data can be processed by rough Sugar Set Theory and wavelet theory.

The dimensions of the data are processed uniformly, and the values of load influencing factors are normalized to the $[0, 1]$ interval; that is,

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_i - x_{\max}}, \quad (2)$$

where x_i and \bar{x}_i are normalized values before and after normalization; x_{\max} and x_{\min} are the maximum value and minimum value in the data sequence, respectively.

3.3. Load Data Partition. The parallel detection algorithm of abnormal data based on Spark density peak clustering can divide the data space into spatial grids. After division, when calculating the local density of data objects, only the sample data objects in the grid cells and the sample data objects in adjacent grid cells need to be considered, which greatly reduces the time complexity of the algorithm [24].

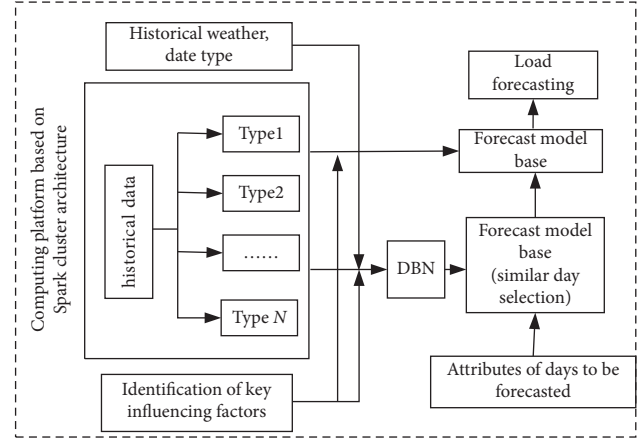


FIGURE 1: Load forecasting process based on improved deep learning in spark architecture method.

In order to divide the data evenly and with the load of each computing node being relatively balanced, to better detect abnormal data, the parallel abnormal data detection algorithm based on Spark density peak clustering uses K-dimensional tree (KD-tree) algorithm to divide the data space into multiple grid cells with roughly the same number of data objects.

When partitioning a data set, some data objects need to be assigned to multiple different partitions at the same time. This is because for the critical points in the grid cell after partitioning and because some adjacent data points within the neighborhood of their density intercept are not in the grid cell but in the adjacent grid cell, if their local density ρ_i is calculated directly. The local density error will be too large, resulting in too large error in anomaly detection. Therefore, in order to calculate the local density ρ_i of these critical points, some data objects need to be allocated to multiple different partitions at the same time [25].

Because the data partition and grid cell are one-to-one correspondence, each data partition corresponds to a grid cell, and each grid cell corresponds to a data partition. In a data partition, the local density and minimum distance of any data object in the data partition can be calculated. The pseudocode of the data partition algorithm is shown in Algorithm 1.

3.4. Abnormal Data Detection. Abnormal data detection is carried out in each partition to eliminate bad data and reduce the impact of data on load forecasting. In the abnormal data detection, firstly set the abnormal value judgment rules and then carry out local clustering anomaly detection in different partitions and merge the local clustering and abnormal points. Finally, the Spark parallel programming model is used to realize the parallel detection of abnormal data.

3.4.1. Abnormal Value Judgment Rules. Based on the density outlier detection method, it is considered that the cluster density of normal sample points is higher than that of outlier

Parameter implication:

Input: X is the data set; n_{\max} is the maximum number of sample data objects in the grid cell.

Output: Partitions is the data partition obtained after the data set is divided.

Begin

(1) Obtain multidimensional data space D_S through sample data set X .

(2) The KD-tree algorithm is used to divide the multidimensional data space D_S into multiple grid cells with relatively balanced size and no coincidence.

(3) The sample data objects are allocated to grid cells, and then the number of sample data objects contained in each grid cell is calculated.

(4) Initialize an empty Queue, add data space D_S to the Queue, and initialize an empty grid cell set D .

(5) Pop up the Queue header element S from the queue, and calculate the number n of sample data objects contained in S .

(6) **If** $n < n_{\max}$, **then**

Add S to D ;

(7) **If** $n \geq n_{\max}$, **then**

Calculate the variance of each dimension of the data object in the m -dimensional space in the spatial area S , select the dimension with the largest variance as the segmentation dimension, divide S into two subspace areas $S1$ and $S2$ with an equal number of data objects, and then add $S1$ and $S2$ to the queue to wait for further division.

(8) **If** Queue is empty, **then**

The result of spatial meshing of data set X is set D ;

Otherwise Skip to step 5.

(9) According to the divided spatial grid set D , all data partitions of the data set can be obtained.

End

ALGORITHM 1: Pseudocode of data partition algorithm.

samples [26]. The proposed method combines the local outlier factor (LOF) algorithm with the density peak clustering algorithm for outlier detection. The specific formula is as follows:

$$\rho_a = \begin{cases} 0, & \text{LOF}_k(a) > 1, \\ \sum_{\text{dist}(x_a, x_j) \leq \sigma} \exp \left[- \left(\frac{\text{dist}(x_a, x_j)}{\text{dist}_{\text{cut off}}} \right)^2 \right], & \text{LOF}_k(a) \leq 1, \end{cases} \quad (3)$$

where σ is the density intercept, and the range less than σ from the data object becomes the density intercept neighborhood of the data object; $\text{LOF}_k(a)$ represents the mean value of the local reachable density ratio of the neighborhood point of point a to point a ; if this ratio is closer to 1, it indicates that the difference between the local reachable density of a and its neighborhood point density is small; $\text{dist}(x_a, x_j)$ is the reachable distance from x_a to x_j ; $\text{dist}_{\text{cut off}}$ is the intercept.

The mathematical description of determining that the sample point is an abnormal sample is as follows:

$$\begin{cases} \rho_a = 0, \\ \delta_a > \delta_{\Theta}, \end{cases} \quad \delta_a + \gamma_a, \quad (4)$$

$$\delta_{\Theta} = \frac{1}{N} \sum_{a=1}^N$$

where δ_a and δ_{Θ} are the relative distance and its threshold, respectively; γ_a is the empirical parameter; N is the total number of samples.

3.4.2. Local Clustering Anomaly Detection in Partition and Outlier Merging. In order to enable each computing node to perform local clustering anomaly detection on its corresponding data partition in parallel, it is necessary to optimize the original density peak clustering. In order to get rid of the intervention of subjective human factors, the original density peak clustering algorithm uses an auxiliary function to select the cluster center. The mathematical expression is as follows:

$$\gamma_i = \rho_i * \hat{\delta}_i, \quad (5)$$

where ρ_i is the local density of sample points and $\hat{\delta}_i$ is the minimum distance of samples.

For local clustering in the partition, a cluster center threshold needs to be given, and the γ value of each sample data object in the data partition is compared with the given cluster center threshold. If the γ value of the sample data object is greater than the set threshold, the data object is regarded as the candidate object of the cluster center.

In order to achieve the goal that each computing node can independently carry out clustering anomaly detection on the corresponding sample data partition, the data partition stage divides the sample data set into several overlapping data partitions, which also contain some common sample data objects. In the stage of local outlier merging and local cluster merging, the algorithm can find out all local clusters to be merged by evaluating the characteristics of these common data objects (i.e., critical points and expansion points). If the outlier sample points repeatedly appear in two or more data partitions, only one outlier sample point needs to be retained to eliminate the duplicate outlier sample points; form a set of global abnormal sample points.

3.4.3. Parallelization of Anomaly Data Detection Algorithm Based on Density Peak Clustering. In the case of massive power load data, the single-machine version of abnormal data detection algorithm is inefficient and cannot meet the requirements of abnormal data detection in power system. Therefore, the Spark parallel programming model is used to parallelize anomaly data detection algorithm to improve its efficiency.

The parallel detection algorithm of density peak clustering anomaly data based on spark mainly includes three important stages, namely, data partition and local clustering in the partition, anomaly detection and local clustering, and anomaly point merging. The algorithm has a lot of distance and density calculation, connection operation, and low efficiency. The execution sequence of the single-machine version of the algorithm can continue to the next stage only after the operation of the previous stage is completed. For this purpose, a clustering algorithm based on peak density is proposed, as shown in Figure 2.

In the Map stage, firstly, the power load data set is divided into multiple grid cells with approximately the same number of data objects by KD-tree algorithm, and then the data partition and grid cells are allocated one by one by data partition algorithm. In the Combine stage, the local clustering anomaly detection algorithm is executed in each partition to obtain the local clustering results of the data partition and the abnormal sample set in the data partition. In the Reduce phase, local cluster merging and outlier merging algorithms are implemented to connect local cluster markers to obtain the clustering results of global clusters and the global set of outlier sample points.

3.5. Load Forecasting Based on Improved Deep Learning in Spark Architecture. In load forecasting, the training data mainly includes load data and meteorological data [27, 28]. For traditional shallow learning methods, on the one hand, due to the simple structure model, it is difficult to learn the complex nonlinear mapping relationship in the training data. On the other hand, in order to avoid local minima, only a small number of features can be selected for training, which fundamentally limits its application scope and prediction accuracy.

The deep learning method based on Spark memory computing framework fully considers the selection of more dimensions and more extensive features, makes maximum use of massive data, and forecasts the load more accurately and quickly. At the same time, considering the characteristics of periodicity and regularity of power load itself, first find the commonness in massive load data by mining the typical load curve of historical load, then classify the load curve categories of forecast days through the load classifier based on DBN, and finally apply the corresponding typical load curve as a feature to the load predictor based on EMD-GRU. The flow chart of load forecasting based on improved deep learning under spark platform is shown in Figure 3.

3.5.1. Historical Load Clustering Based on Spark. After clustering the load curve by fuzzy C-means clustering (FCMC), the load characteristics of n distribution

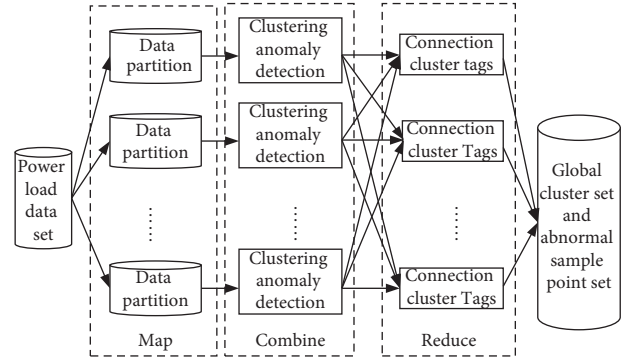


FIGURE 2: Algorithmic parallel framework.

transformer in the distribution network are classified into c homogeneous clusters. In order to solve the optimal membership matrix U and clustering center matrix Ψ , the following objective functions can be constructed according to the clustering criteria:

$$J(U, \Psi, \lambda) = \sum_{i=1}^n \sum_{k=1}^c (u_i^k)^\tau (\text{dist}_i^k)^2 + \sum_{i=1}^n \lambda_i \left(\sum_{k=1}^c u_i^k - 1 \right), \quad (6)$$

where $u_i^k \in [0, 1]$ represents the degree that the i -th distribution transformer belongs to the k -th cluster center, and the sum of membership degrees of one distribution transformer to all clusters is equal to 1; $(\text{dist}_i^k)^2$ is the Euclidean distance between the i -th distribution transformer and the k -th cluster center; $\tau \in [0, 2]$ is the weighted index; λ_i is the Lagrange multiplier of equality constraint. The iterative formula to minimize the objective function is calculated as follows:

$$u_i^k = \frac{1}{\sum_{q=1}^c (\text{dist}_i^k / d_i^q)^{2/(\tau-1)}}, \quad (7)$$

$$\varphi_k = \frac{\sum_{i=1}^n (u_i^k)^\tau x_i}{\sum_{i=1}^n (u_i^k)^\tau},$$

where x_i is the data volume of the i -th distribution transformer; φ_k is the cluster center quantity of fuzzy clustering.

In order to maintain the consistency and rapidity of the whole load forecasting process, the fuzzy C-means clustering algorithm based on spark memory computing environment is used. The specific steps include the following:

- (1) *Drive*. The main task is to initialize the basic functions of the program and drive each subtask through the function method in Spark. After the cluster is started, each node in the cluster will load each row of data in the data set file into Spark as a resilient distributed data set (RDD) and copy the shared data to each node in the cluster.
- (2) *Map Task*. The RDD of the load data set is calculated by line-by-line scanning, and the distance between the data object and each central point in the global variable cluster center is calculated. By comparison,

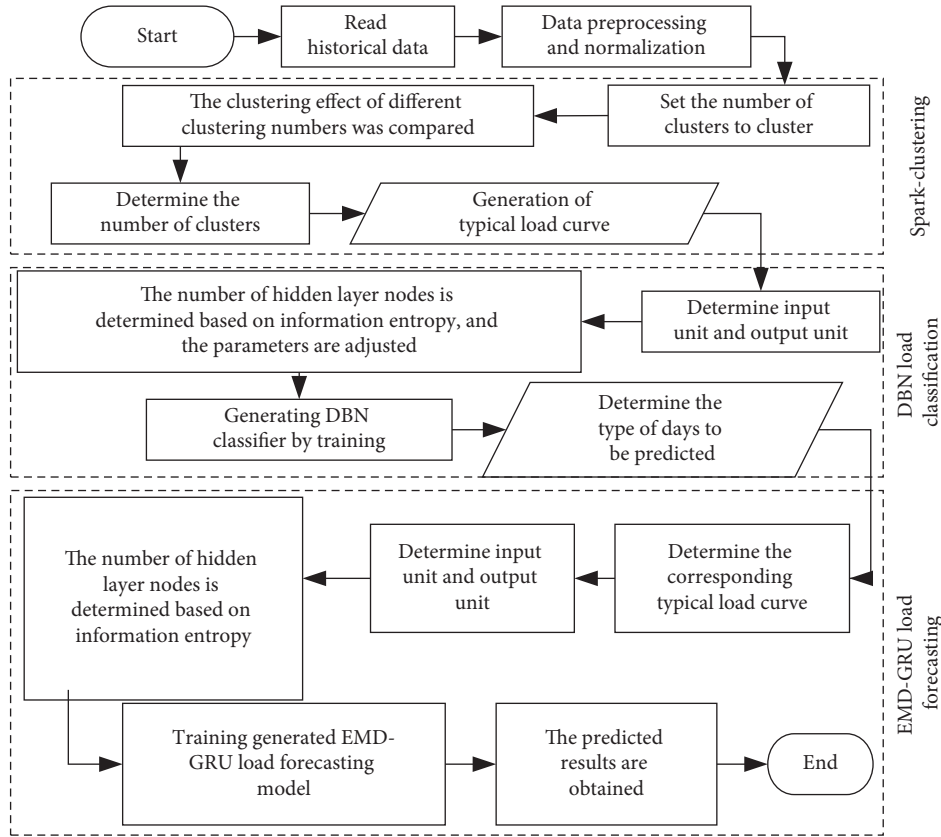


FIGURE 3: Overall framework.

the data object is assigned to the cluster with the smallest distance from the central point, and the final output is the key value pair $\langle \text{key}, \text{value} \rangle$, where key is the cluster center and value is the data object belonging to the cluster center.

- (3) *Combine Task.* After the data set is mapped, a large number of RDD intermediate data sets will be generated. In order not to make the network communication a bottleneck, the values belonging to the same key are averaged locally, the local results $\langle \text{key}, \text{value} \rangle$ are obtained, and then the data are transmitted to the master node for processing to reduce the traffic.
- (4) *Reduce Task.* Summarize and merge the local results of the Combine process from the calculation node and return the result RDD in the form of an array to generate a global result. The number of Combine data points in each calculation node is different. The counter is used to count the data points to obtain a weight. During Reduce calculation, the weight and local results are used to calculate the global result.

3.5.2. Classification of Daily Load Types to be Forecasted Based on DBN. In order to obtain more accurate load forecasting results, the load categories on the forecast day shall be classified before the actual load forecasting. Then, the typical load curve of the corresponding category is

extracted, and the load in the curve is taken as a relevant input, so as to reflect the hidden commonness of massive load in the process of load forecasting. Because this process is essentially equivalent to extracting the characteristics of historical load and classifying the typical load curve and the probability generation model DBN formed by the stacking of restricted Boltzmann machine (RBM) has very strong learning ability in classification, the proposed method selects DBN to determine the load category of the day to be predicted.

The specific steps of classifying forecast days based on DBN are as follows:

- (1) Determine the input element and output element of DBN. Generally, the selected input elements include influencing factors related to the daily load to be predicted, such as load data on the same regular date as the day to be predicted, as well as its category, meteorological data, date attribute, and so forth. For the date attribute in the input element, convert it to 8 binary inputs. The setting of output element depends on the result of FCMC.
- (2) Based on information entropy theory, the number of neurons in the hidden layer of DBN is determined.
- (3) Stack RBM to form DBN and adjust the relevant parameters of the network. The parameter adjustment is mainly obtained through the continuous verification and analysis of the model effect.

- (4) Save the training data to HDFS and convert it to RDD format.
- (5) The DBN model is trained in parallel through data parallel. The specific method is to establish multiple data slices in spark cluster, create multiple copies of neural network model, train each slice at the same time, and cache the intermediate results and extract the training speed of the model from memory. After the training of each copy, the calculated parameter adjustment value is transmitted to the model parameter server, and a new parameter is applied to the parameter server for the next step of training.
- (6) The trained DBN model is used to determine the load curve category of the day to be predicted.

3.5.3. Load Forecasting Based on EMD-GRU. There will be some high-frequency noise components in the load subsequence based on EMD decomposition, which will affect the overall prediction accuracy. Therefore, an EMD-GRU load forecasting model based on feature selection is proposed. Feature selection of decomposed subsequences can not only avoid multiple prediction errors and improve prediction accuracy but also reduce prediction workload and model complexity.

The overall framework of load forecasting based on EMD-GRU is shown in Figure 4.

Firstly, the raw load sequence is decomposed by EMD into intrinsic mode functions (IMF) and residuals containing different characteristics of the original load series. Then, the original feature set is analyzed and screened by Pearson correlation coefficient method. Finally, combined with the selected time series characteristics and the raw load sequence, it is input into GRU model to realize power grid load forecasting. The Pearson correlation coefficient method is calculated as follows:

$$\eta = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (8)$$

where η is the correlation coefficient, x_i and y_i are the sample points, \bar{x} and \bar{y} are the sample means, and n is the number of samples.

4. Experiment and Analysis

4.1. Experimental Environment and Data Set. The Spark cluster built in the experiment is composed of 8 PCs with the same configuration. Each PC has 8G memory, 2T hard disk, and dual-core Intel i7 CPU, the main frequency is 4.7 GHz, and it runs Centos Linux operating system. One machine is the Master node, which is responsible for resource allocation and job scheduling of the whole cluster, and the other 7 machines are slave nodes, which are mainly used to store data and run tasks. The Spark cluster topology is shown in Figure 5.

The experimental data comes from the load data and influencing factor data collected by a regional power grid. The amount of data is TB and it has high dimensions. It is

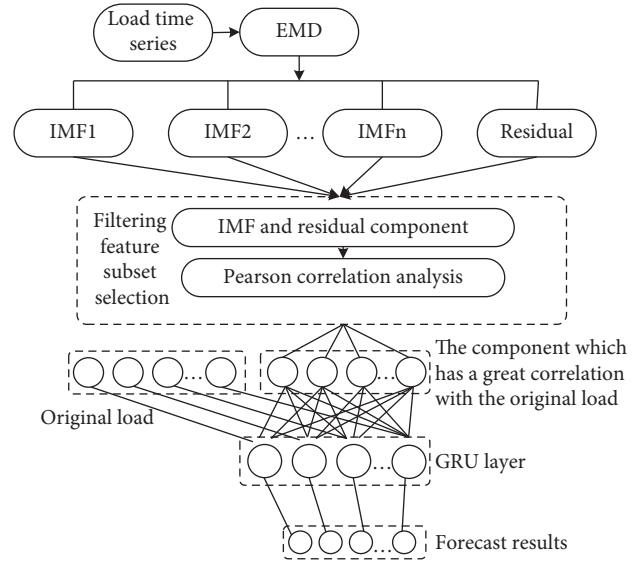


FIGURE 4: Overall block diagram of load forecasting based on EMD-GRU.

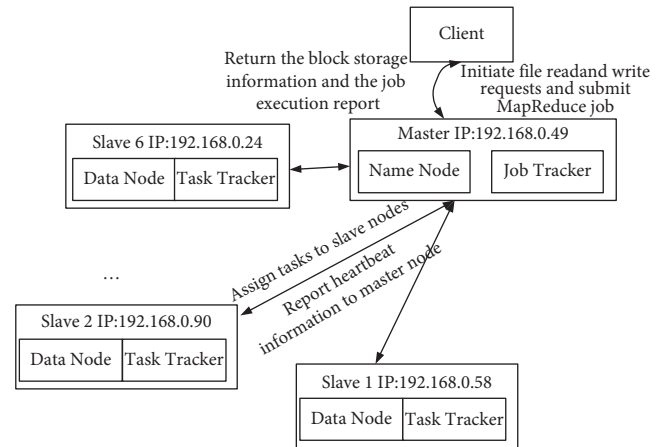


FIGURE 5: Topology of Spark cluster.

mainly structured and semistructured data, which is in line with the characteristics of power big data. The training sample is the power consumption data from May 1, 2020, to August 25, 2020, and the sampling interval is 1 h. Taking the 24-hour load data from May 1 to May 25, 2021, as the test sample, the load forecasting effect is evaluated by relative error, average error E_{ME} , root mean square error E_{RMSE} , and average absolute percentage error E_{MAPE} . The evaluation indexes are calculated as follows:

$$e_{ME} = \frac{1}{n+1} \sum_{i=0}^n (y(i) - \hat{y}(i)),$$

$$e_{RMSE} = \sqrt{\frac{\sum_{i=0}^n (y(i) - \hat{y}(i))^2}{n+1}}, \quad (9)$$

$$e_{MAPE} = \frac{1}{n+1} \sum_{i=0}^n \left| \frac{y(i) - \hat{y}(i)}{\hat{y}(i)} \right| \times 100\%,$$

where $y(i)$ and $\hat{y}(i)$ are the true load value and predicted value at time i .

4.2. Effectiveness Experiment. In order to demonstrate the effectiveness of the proposed method, the proposed method and the method in [22] are used to predict the power load, respectively. The load forecasting results in a short time are shown in Figure 6.

As can be seen from Figure 6 that the predicted value of the proposed method is closer to the real value, while the prediction deviation of the method in [22] is obvious, in order to quantitatively analyze the accuracy of load forecasting by the proposed method, six time points are selected for experimental comparison, and the results are shown in Table 1.

It can be seen from Table 1 that the accuracy of the proposed method has reached a high level, and the average prediction error is basically controlled within 3%, while the maximum relative error of the method in [22] is 6.35%.

4.3. Prediction Accuracy Analysis. In order to better demonstrate the load forecasting accuracy of the proposed method in a long time, the load within 10 days is predicted. The results of the proposed method and the load forecasting error evaluation index in [22] are shown in Table 2. Among them, January 1-3 is the new year's day holiday, January 4-8 are the working days, and January 9-10 are the weekend, which can better reflect the prediction effect under each load mode.

As can be seen from Table 2 that, in long-term load forecasting, the error evaluation index value of the proposed method is better than the comparison method. e_{ME} , e_{RMSE} , and e_{MAPE} are 0.023 MW, 19.75%, and 2.74% respectively. The mean value meets the assessment index of the State Grid, and the overall deviation control is better. Because the proposed method uses the density peak clustering algorithm to detect abnormal data and improve the quality of data set and uses the combination of DBN and EMD-GRU to realize load forecasting, it can ensure high forecasting accuracy. However, the model in [22] is complex and the control of data quality is insufficient, so the overall forecasting accuracy is lacking. In addition, the prediction accuracy of the proposed method in some time periods is not ideal, but the prediction errors are within the national grid standard.

In addition, the load forecasting error curve of the proposed method and the methods in [16], [18], and [22] within 48 h is shown in Figure 7.

As can be seen from Figure 7, the prediction error of the proposed method within 24 h is less than that of other comparison methods, with a minimum of about 1.31% and a maximum of about 3.38%. Except for individual points, the prediction error at most times is less than 3%. The proposed method ensures the accuracy of load data through data preprocessing, data partition detection, and other operations. On the basis of load classification by DBN, EMD-GRU model is used to reduce the load forecasting error to a great extent. Reference [22] was based on the method of using stacking condition limited Boltzmann machine to realize fast

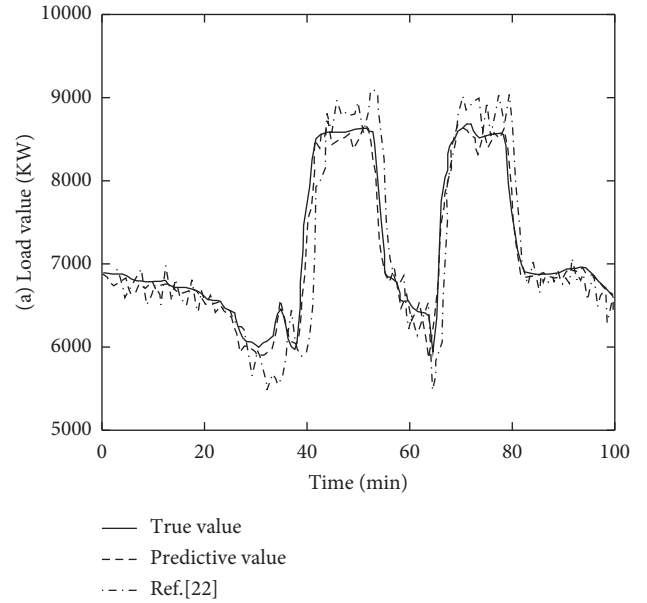


FIGURE 6: Load forecasting results of different methods in short time.

TABLE 1: Comparison results of different load forecasting methods.

No.	True value (MW)	Predictive value (MW)	
		Ref. [22]	The proposed method
1	8.254	9.106	7.922
2	8.739	9.054	8.623
3	7.825	8.234	7.711
4	7.241	8.338	7.489
5	8.632	9.104	8.532
6	8.963	9.272	8.645

and accurate short-term load forecasting, but it performed poorly in long-term load forecasting, and the growth rate of prediction error in the later stage was obvious. Reference [18] combined the least squares support vector machine learning engine and the improved virus colony search algorithm to realize load forecasting and used the wavelet transform method to extract data features. The prediction error is small in a short time, and the prediction error increases with the passage of time, due to its lack of full consideration of influencing factors. Reference [16] uses sequence to sequence cyclic neural network with attention to predicting power load. The model is single, the overall performance is poor, and it depends too much on time attributes, up to 4.68%.

4.4. Parallel Performance Analysis. With the increase of the amount of input data, the traditional serial processing method is difficult to meet the requirements of load forecasting. The parallel algorithm based on Spark can calculate the whole forecasting task in parallel to improve the computing efficiency. The speedup ratio is an important standard to measure the parallel efficiency of a parallelized system. Amdahl's law points out that the better the parallel

TABLE 2: Comparison results of different load forecasting methods.

Date	e_{ME} (MW)		e_{RMSE} (%)		e_{MAPE} (%)	
	The proposed method	Ref. [22]	The proposed method	Ref. [22]	The proposed method	Ref. [22]
1	-0.172	-0.279	19.73	22.57	3.41	4.25
2	0.357	0.750	21.14	20.62	3.25	3.97
3	0.136	0.621	16.75	16.30	2.59	2.45
4	0.173	0.237	23.57	25.79	2.77	3.72
5	-0.437	-0.324	21.49	25.41	2.09	3.34
6	0.574	-0.349	17.74	24.75	2.33	3.47
7	-0.275	0.427	19.27	19.54	2.74	2.95
8	-0.234	-0.732	17.54	22.73	3.12	3.57
9	0.314	0.449	22.73	21.74	2.47	2.75
10	-0.205	0.407	14.15	19.35	2.59	2.99
Average value	0.023	0.139	19.75	21.90	2.74	3.37

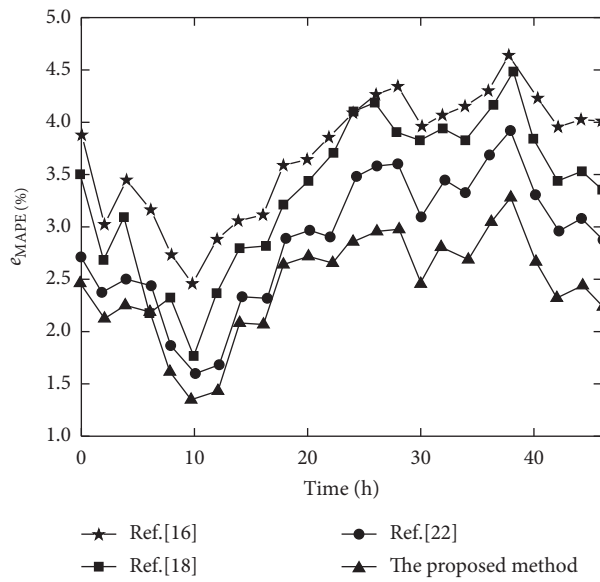


FIGURE 7: Comparison of load forecasting errors of different methods.

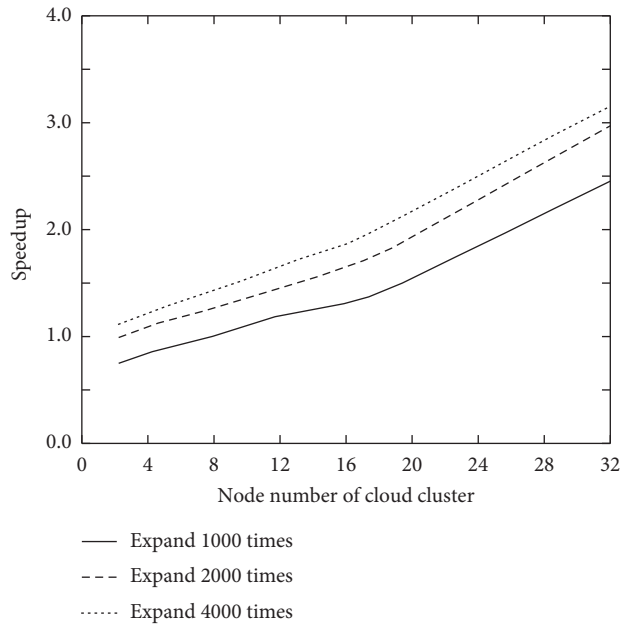


FIGURE 8: Parallelization performance results.

computing system, the speedup ratio is closer to the number of processors in the parallelized system. Therefore, in the parallel performance experiment, when the number of cluster nodes changes from 2 to 32, the acceleration ratio results of the cloud computing platform are shown in Figure 8.

As can be seen from Figure 8, when there are more than 16 cloud cluster nodes, additional consumption such as network transmission between nodes will increase, so the acceleration ratio will deteriorate with the increase of cloud cluster nodes. However, with the increase of the amount of data, the speedup ratio of this method still increases almost linearly, indicating good parallel performance.

5. Conclusions

The load forecasting method based on improved deep learning in cloud computing environment is proposed in this paper. The experimental results show that this method has high prediction accuracy and good parallel performance. However, in load forecasting based on deep learning algorithm, the proposed method still relies on manual selection when selecting the relevant features of load forecasting and does not make full use of the feature extraction ability of unsupervised learning in deep learning. In the next research, we can study how to automatically extract the relevant features of load forecasting.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Q. Wang and F. Jiang, "Integrating linear and nonlinear forecasting techniques based on grey theory and artificial intelligence to forecast shale gas monthly production in Pennsylvania and Texas of the United States," *Energy*, vol. 178, no. 6, pp. 781–803, 2019.
- [2] S. Sengar and X. Liu, "An efficient load forecasting in predictive control strategy using hybrid neural network," *Journal of Circuits, Systems, and Computers*, vol. 29, no. 1, pp. 567–576, 2020.
- [3] D.-J. Bae, B.-S. Kwon, C.-H. Moon, S.-H. Woo, and K.-B. Song, "Short-term load forecasting algorithm on weekdays considering the amount of behind-the-meter generation," *Journal of the Korean Institute of Illuminating and Electrical Installation Engineers*, vol. 34, no. 11, pp. 37–43, 2020.
- [4] S. O. Khomutov and N. A. Serebryakov, "Creation of a short-term load forecasting neural network model of electrical engineering complex of section regional electric grid 6-35 kV," *Transportation Systems and Technology*, vol. 6, no. 1, pp. 80–91, 2020.
- [5] G. Fernandez, V. Krishnasamy, J. Ali, J. Sathik, and S. Aleem, "Internet of things based real-time electric vehicle load forecasting and charging station recommendation," *ISA Transactions*, vol. 97, no. 4, pp. 431–447, 2020.
- [6] D. Ma, "Short-term load forecasting based on variable weighted synthesis of different kernel SVM," *Statistics and Applications*, vol. 9, no. 1, pp. 73–80, 2020.
- [7] P. D. Lê, D. M. Bui, D. A. Phm et al., "Applying statistical analysis for assessing the reliability of input data to improve the quality of short-term load forecasting for a Ho Chi Minh City distribution network," *Science & Technology Development Journal - Engineering and Technology*, vol. 2, no. 4, pp. 223–239, 2020.
- [8] A. Io, B. Sbe, and B. Ho, "A combined deep learning application for short term load forecasting," *Alexandria Engineering Journal*, vol. 60, no. 4, pp. 3807–3818, 2021.
- [9] M. Sun, T. Zhang, Y. Wang, G. Strbac, and C. Kang, "Using bayesian deep learning to capture uncertainty for residential net load forecasting," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 188–201, 2020.
- [10] F. Z. Abera and V. Khedkar, "Machine learning approach electric appliance consumption and peak demand forecasting of residential customers using smart meter data," *Wireless Personal Communications*, vol. 111, no. 1, pp. 65–82, 2020.
- [11] Q. Chen, M. Xia, T. Lu, X. Jiang, W. Liu, and Q. Sun, "Short-term load forecasting based on deep learning for end-user transformer subject to volatile electric heating loads," *IEEE Access*, vol. 7, pp. 162697–162707, 2019.
- [12] L. Wen, K. Zhou, S. Yang, and X. Lu, "Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting," *Energy*, vol. 171, pp. 1053–1065, 2019.
- [13] D. Kim, H. J. Jo, M. S. Kim, J. H. Roh, and J.-B. Park, "Short-term load forecasting based on deep learning model," *The Transactions of the Korean Institute of Electrical Engineers*, vol. 68, no. 9, pp. 1094–1099, 2019.
- [14] J. Ye, Y. Dang, and Y. Yang, "Forecasting the multifactorial interval grey number sequences using grey relational model and GM (1, N) model based on effective information transformation," *Soft Computing*, vol. 24, no. 7, pp. 5255–5269, 2020.
- [15] G. Suryanarayana, J. Lago, D. Geysen, P. Aleksiejuk, and C. Johansson, "Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods," *Energy*, vol. 15, no. 3, pp. 141–149, 2018.
- [16] L. Sehovac and K. Grolinger, "Deep learning for load forecasting: sequence to sequence recurrent neural networks with attention," *IEEE Access*, vol. 8, pp. 36411–36426, 2020.
- [17] A. Ghasemi-Marzbali, "A developed short-term electricity price and load forecasting method based on data processing, support vector machine, and virus colony search," *Energy Efficiency*, vol. 13, no. 7, pp. 1–18, 2020.
- [18] H. J. Sadaei, P. C. de Lima e Silva, F. G. Guimarães, and M. H. Lee, "Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series," *Energy*, vol. 175, no. 5, pp. 365–377, 2019.
- [19] Z. Wang, B. Zhao, and W. Ji, "Short-term load forecasting method based on GRU-NN model," *Dianli Xitong Zidonghua/Automation of Electric Power Systems*, vol. 43, no. 5, pp. 53–58, 2019.
- [20] M. Cai, M. Pipattanasomporn, and S. Rahman, "Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques," *Applied Energy*, vol. 236, pp. 1078–1088, 2019.
- [21] F. Mehmood Butt, L. Hussain, L. Hussain, A. Mahmood, and K. Javed Lone, "Artificial Intelligence based accurately load

- forecasting system to forecast short and medium-term load demands,” *Mathematical Biosciences and Engineering*, vol. 18, no. 1, pp. 400–425, 2021.
- [22] L. Barolli, F. Xh Afa, and N. Javaid, “Short term load forecasting based on deep learning for smart grid applications,” *Advances in Intelligent Systems and Computing*, vol. 10, no. 25, pp. 276–288, 2019.
- [23] S. Smyl, N. G. Hua, and R. J. Hyndman, “Machine learning methods for GEFCom2017 probabilistic load forecasting,” *International Journal of Forecasting*, vol. 35, no. 4, pp. 1424–1431, 2019.
- [24] G. Fu, “Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system,” *Energy*, vol. 148, pp. 269–282, 2018.
- [25] J. Wang, X. Chen, F. Zhang, F. Chen, and Y. Xin, “Building load forecasting using deep neural network with efficient feature fusion,” *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 1, pp. 160–169, 2021.
- [26] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, “Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches †,” *Energies*, vol. 11, no. 7, pp. 1636–1645, 2018.
- [27] Z. Kong, C. Zhang, H. Lv, F. Xiong, and Z. Fu, “Multimodal feature extraction and fusion deep neural networks for short-term load forecasting,” *IEEE Access*, vol. 8, no. 3, pp. 185373–185383, 2020.
- [28] W. Kim, Y. Han, K. J. Kim, and K.-W. Song, “Electricity load forecasting using advanced feature selection and optimal deep learning model for the variable refrigerant flow systems,” *Energy Report*, vol. 6, no. 2, pp. 2604–2618, 2020.

Research Article

Research and Design of Distributed Fire Alarm System of Indoor Internet of Things Based on LoRa

Wei Chen,¹ ChenYu He,² JianRong Lu ,³ Kui Yan,¹ Jin Liu,¹ Feng Zhou,¹ Xin Xu,² and Xiao Hao²

¹Industrial Center/School of Innovation and Entrepreneurship, Nanjing Institute of Technology, Nanjing, Jiangsu 211100, China

²Graduate School, Nanjing Institute of Technology, Nanjing, Jiangsu 211100, China

³Aviation Engineering Institute, Jiangsu Aviation Technical College, Zhenjiang, Jiangsu 211100, China

Correspondence should be addressed to JianRong Lu; lujianrong@jatc.edu.cn

Received 6 August 2021; Revised 1 September 2021; Accepted 6 September 2021; Published 8 October 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Wei Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to comprehensively improve the sensitivity of fire warning and effectively shorten the warning time, this paper proposes and implements an indoor distributed fire alarm system based on low power wide area network. The system is mainly composed of three parts: a multisensor acquisition node based on LoRa technology, a distributed edge gateway, and a remote user monitoring system. The multisensor collection node obtains environmental parameters such as indoor temperature, smoke concentration, and air quality and then transmits the sensing data to edge gateway by LoRa after preprocessing. The edge gateway is based on an embedded Linux platform and is deployed in distributed state to collect and store data from multiple collection nodes. Besides, edge gateway forwards valid data to the remote user monitoring system by standard MQTT protocol. The user monitoring system displays current deployment area parameters to users in real time and provides early warning prompts based on relevant preset indicators to help the administrator make more accurate decisions on corresponding measures. The system has been deployed and tested in Nanjing Institute of Technology. By sensor calibration experiments, LoRa communication experiments, and system tests in different environments, the experimental results show that the average received signal strength in a small interference space is -104.12 dBm, and the average received signal strength in a noisy signal environment is -57.5 dBm. By setting the optimal transmitting power for each distance, the packet receiving rate can reach more than 95%, and the alarm accuracy can reach 100% under premise of ensuring the lowest power consumption. Finally, this paper conducts a comprehensive performance analysis on the wireless communication performance of environmental collection nodes, multisensor data fusion algorithm, distributed LoRa edge gateway deployment performance, and remote system early warning accuracy.

1. Introduction

Nowadays, fire is a common serious problem faced by countries all over the world. While causing casualties, it will also bring certain property losses and adverse sociopolitical effects. According to statistics from the China Commercial Fire and Safety Association in 2019, a total of 233,000 fires were reported across the country, 1,335 people died, and direct property losses reached 3.612 billion yuan. Compared with 2018, the numbers of fires and property losses have been reduced by 4% and 1.9%, respectively [1]. Among them, the number of fires caused by electrical short-circuit aging and other reasons is majority. Figure 1 shows the situation of

serious fires in the past 13 years. It can be seen from the figure that the number of fires is decreasing year by year [2]. In addition, with the continuous growth of population and economy, high-rise buildings are also increasing, and the number of fires has increased by 5.8% compared with the previous year. Therefore, high-rise buildings have become the focus and difficulty of fire prevention. The pressure and difficulty of urban fire prevention and control will continue to increase, and risks must be predicted in advance and preventive measures must be studied in advance. Detecting the occurrence of fire in the first time and monitoring fire in real time can play a vital role in reducing the loss of life and property.

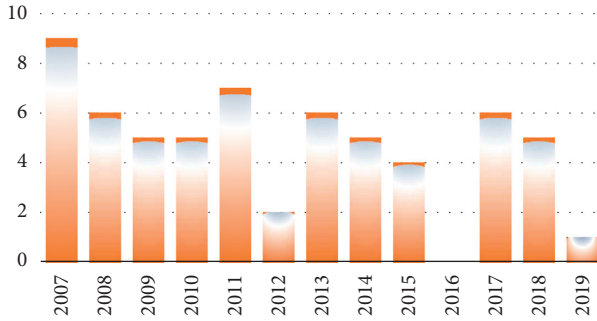


FIGURE 1: The number of extreme fires in 13 years.

With the development of Internet of Things (IoT) technology, the application of IoT technology to fire alarms can accurately locate the fire point and has the characteristics of real-time monitoring, fire tracking, online upgrades, and convenient maintenance. Many countries have also conducted a lot of research on wireless communication technology applications to indoor fire protection. For example, scholars such as Bo Chang studied the indoor fire monitoring system that used ZigBee wireless network to monitor the temperature, humidity, and smoke concentration of fire site in real time [3]. Liu Yunhong and other scholars studied the use of ZigBee to Wi-Fi communication gateway to collect the data collected by sensor nodes to gateway and upload it to servers to solve the indoor location problems of fire locations and evacuation plans [4]. Scholars such as SwetaBasu adopted Wireless Sensor Network (WSN) to accumulate accurate environmental data by sensor nodes such as temperature, relative humidity, and different gas concentrations and send these data to base stations connected to the ground [5, 6]. In 2019, Zhou et al. [7] used an infrared thermal imager to sense changes in indoor temperature and sent alarm messages by 4G module to remind users by sending text messages. However, the data packets in mobile network need to be uploaded by the base station of operator, [8] which increases a large part of communication cost. ZigBee wireless network technology has the characteristics of low complexity, low power consumption, and low cost [9]. ZigBee technology has its own inapplicability in this field due to its short communication distance [10].

Based on the above problems, this paper chooses LoRa as communication module. The module has long distance and low power consumption characteristics. The system is divided into three parts: node, gateway, and cloud server. The nodes and gateways are deployed in a star network structure. As shown in Figure 2, this system is deployed in urban indoor buildings, large warehouses, or large factories and installed in ventilated and easy-to-detect locations such as walls and ceilings. Compared with other solutions, this solution differs in that terminal nodes are easy to deploy. Moreover, the deployed nodes have a long service life and through sensor filtering and weighted fusion algorithm to judge whether a fire occurs. Each node can remotely monitor the site in real time by remote user monitoring system through the web page.

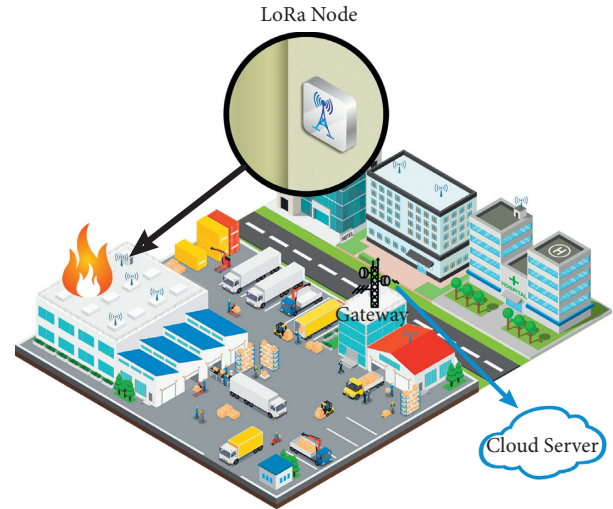


FIGURE 2: Low power alarm system solution.

The rest of this article is organized as follows. Section 2 introduces the system architecture and implementation functions, as well as an overview of LoRa applications. Section 3 introduces the software and hardware design of nodes and gateways. Section 4 discusses the architecture and design of the remote user monitoring system. The performance parameters of the system through the test prototype are calculated in Section 5, which explains the conclusion and direction of future improvement. Finally, this paper is concluded in Section 6.

2. Related Works

2.1. Monitoring System Structure. The system block diagram described in this paper is shown in Figure 3, which is composed of LoRa node, LoRa gateway, cloud server, and remote user monitoring system.

The sensor on LoRa node collects ambient temperature, smoke concentration, and carbon monoxide content at a certain time interval. The collected data is fused and calculated by fire support weighting algorithm, and weighted fusion result is compared with the set threshold. If the fire alarm threshold is reached, an alarm signal is issued and a data packet with a fire alarm signal is sent to the gateway. When there is no fire, the node periodically checks battery power. If battery power is lower than the preset threshold, node's red Light Emitting Diode (LED) will flash and upload a data packet with a low battery alarm signal to the gateway. When the node is not working, it enters a low power consumption mode to extend battery life.

The gateway uses embedded processor IMX6 based on ARM Cortex-A7 core introduced by NXP. It is used to receive data sent from the node. After the data is sorted and packaged, it is sent to cloud server by Transmission Control Protocol (TCP).

The cloud server is mainly composed of Message Queuing Telemetry Transport (MQTT) Broker, MySQL, Backend Service, and Tomcat. When gateway receives data sent from the node, it encapsulates the data and publishes it

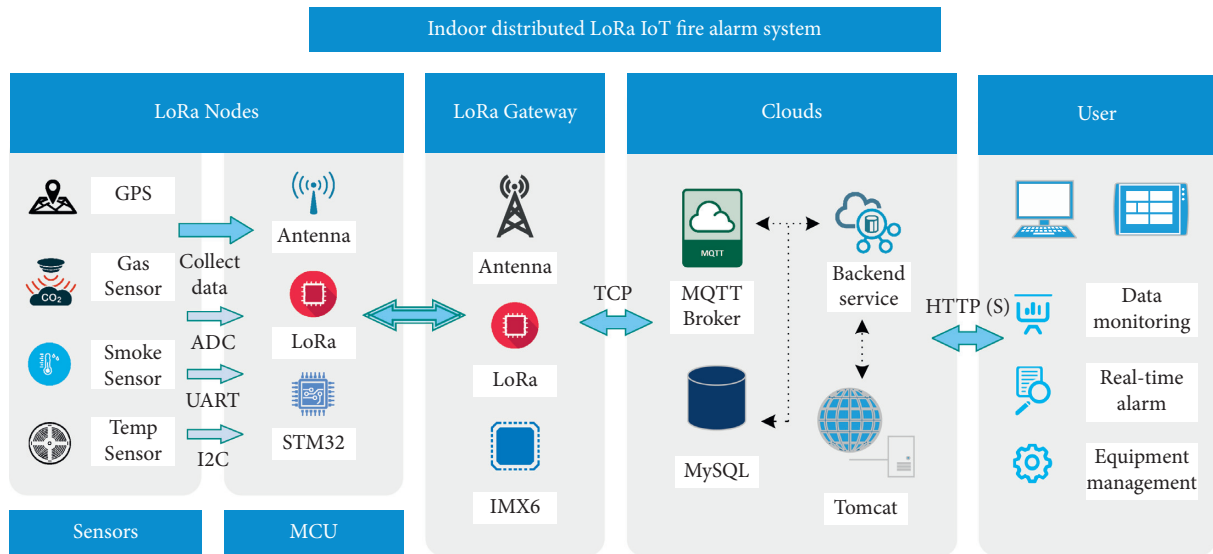


FIGURE 3: System architecture diagram.

under the corresponding topic of MQTT Broker. Then all remote user monitoring systems (MySQL, RESTful Web Server) subscribed to this topic will receive messages.

- (1) MQTT Broker is a message forwarding service provided by servers. Client is a remote user monitoring system used to connect to MQTT Broker to subscribe and publish messages. The IMX6 gateway runs TCP-based MQTT protocol software (MQTT Client) and connects to the Broker and creates a theme representing the gateway in MQTT theme. When gateway receives the data, it will publish the data package under this topic. At the same time, by plug-in of MQTT Broker, MySQL database can be subscribed to specific topics on MQTT Broker. When the topic has data released, back up the data to MySQL.
- (2) Backend Service provides a web page that can manage node information, and users access the server's address by a browser. Corresponding data can be displayed on web page, and users can also perform corresponding operations. The Web Server subscribes to messages on MQTT Broker and queries data in the database in background. When necessary, the message can also be published for gateway to receive and transmit to the node.

The remote user monitoring system realizes the visual monitoring of data and can observe the fire situation and battery power information at each node. At the same time, the communication parameters of nodes can be set in the remote user monitoring system (transmission frequency, transmission power, spreading factor, etc.). You can choose to set a single target node, or you can choose the broadcast mode to set all nodes in the range.

2.2. LoRa Technology Overview. LoRa communication can effectively fight against Doppler frequency shift, using Chirp Spread Spectrum (CSS) technology to send data [11, 12].

This technology is based on symmetrical bilateral and dual-channel ranging positioning technology, covering a large distance, which can effectively reduce noise and interference. The low-cost feature also makes it one of the reasons why it is so hot. The average cost of a typical low power wide area network (LPWAN) module is \$4–18 [13], but the price of LoRa chip is about \$1–2, and the price of complete module is about \$5–10. This also makes LoRa technology regarded as the future wireless communication standard of IoT in the industrial field [14, 15].

SX1278 transceiver mainly uses LoRa remote modem. The module has good anti-interference ability and can minimize the current consumption during work. Under the condition of ensuring high sensitivity exceeding -148dBm, constant RF power output can be achieved when the voltage changes. This solves the traditional communication scheme, when long-distance communication cannot guarantee a certain degree of sensitivity, anti-interference, and power consumption [16].

3. Materials and Methods of Data Acquisition Device Terminal

3.1. LoRa Node Hardware Implementation. The node hardware block diagram is shown in the figure. The core unit on the main board for subnode is composed of a control unit, a radio frequency unit, a sensor interface, an acoustooptic indicating unit, a data storage unit, a debugging interface, and a power supply unit. The Microprocessor Control Unit (MCU) uses STM32L151C9T6 chip to control data collection, data storage, and manage the communication between data and gateway. The sub-node prototype and hardware block diagram are shown in Figure 4. For the abovementioned sensors of different communication types, they can be connected to sensors interfaces reserved at both ends of the board. Each node has its own ID number and communicates with the gateway by different channels.

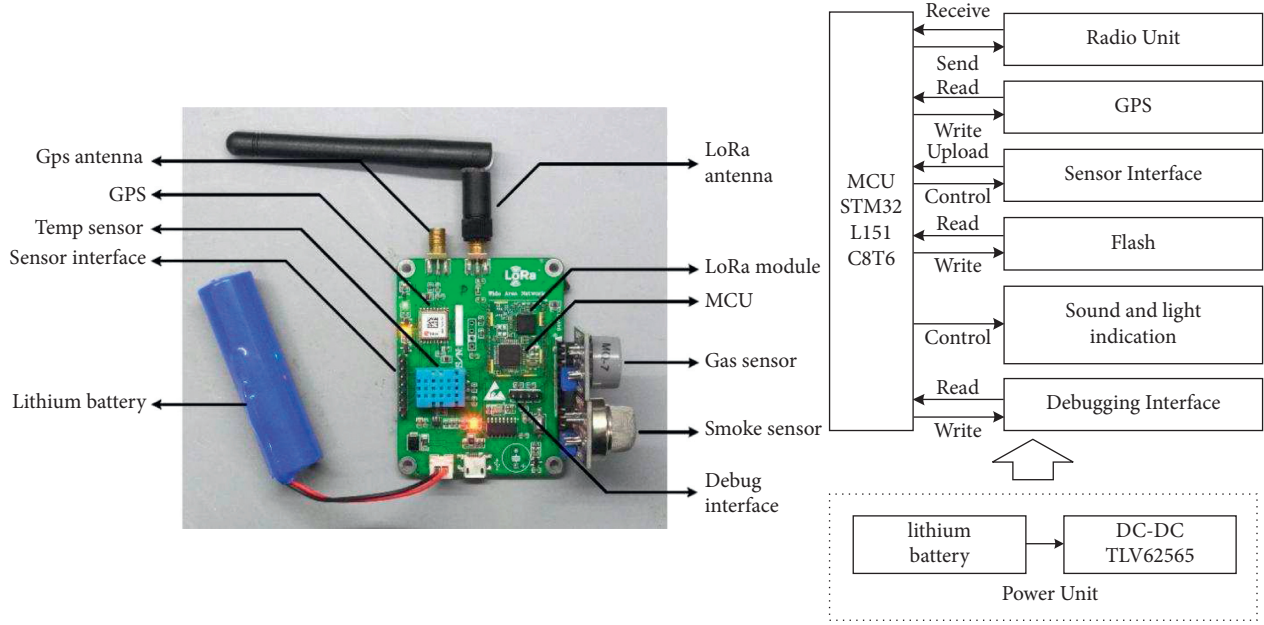


FIGURE 4: Schematic diagram of node hardware.

3.1.1. Hardware Materials. In order to detect whether a fire occurs in real time and monitor the fire situation during a fire, subnode development board records the corresponding data by following sensors: temperature ($^{\circ}$), humidity (%), carbon dioxide level (CO_2 ppm), carbon monoxide level (CO ppm), and smoke concentration. Besides, the remaining power of subnode power supply battery is also monitored, and the node peripherals are shown in Table 1.

The main control unit chip adopts ultra-low power MCU STM32L151 series processor produced by STMicroelectronics. This series of embedded processors uses a high-performance ARM Cortex-M3 32 bit RISC core with a frequency of up to 32 MHz and five different low power modes [17]. In order to reduce costs and the size of subnodes, without wasting hardware resources, various factors are comprehensively considered. This node uses STM32L151C8T6 processor in the series. This model integrates a 12 bit, 14-channel digital-to-analog converter; Three Universals Synchronous/asynchronous serial Receiver/Transmitter (USART) ports, Interintegrated Circuit (I2C) [18]; Two-way Serial Peripheral Interface (SPI) and 37 General-Purpose Input/Output (GPIO) pins and built-in 64 kBytes Flash and 32 kBytes RAM. Use 1.8 to 3.6 V DC power supply for power supply [19]. The processor is widely used in medical equipment, GPS and sports equipment, and wired and wireless sensors. This module is used to control the data acquisition, processing, storage, and control of radio frequency module to send and receive information.

The SX1278 transceiver adopts Semtech's LoRa modulation technology, which can carry out ultra-long-distance spread spectrum communication and has good anti-interference performance [20, 21]. At the same time it minimizes current consumption. Through the integrated +20dBm power amplifier, long-distance wireless communication can be carried out under the condition of ensuring sensitivity as

low as -148dBm, which solves the contradiction between range, anti-interference, and energy consumption.

Use DHT-11 sensor to measure temperature and humidity and use MQ-2 sensor to measure smoke and combustible gas; MQ-7 sensor can detect the concentration of carbon monoxide (CO) [22–24].

3.1.2. Power Module. The power supply unit is composed of a lithium battery and a step-down circuit. The lithium battery provides a DC 5 V working voltage, and 5 V voltage is converted into a suitable working voltage DC 3.3 V for MCU through step-down chip TLV62565. The step-down circuit is shown in Figure 5. Among them, TLV62565 chip cooperates with resistors R23 and R26 to generate a suitable working voltage [25, 26].

$$\begin{aligned}
 V_{\text{OUT}} &= V_{\text{FB}} \times \left(1 + \frac{R_{26}}{R_{23}} \right) \\
 &= 0.6V \times \left(1 + \frac{R_{26}}{R_{23}} \right) = 3.3V, \\
 R_{23} &= \frac{V_{\text{FB}}}{I_{\text{FB}}} = \frac{0.6V}{5\mu A} = 120k\Omega, \\
 R_{26} &= R_{23} \times \left(\frac{V_{\text{OUT}}}{V_{\text{FB}}} - 1 \right) \\
 &= R_{23} \times \left(\frac{V_{\text{OUT}}}{0.6V} - 1 \right) = 27k\Omega.
 \end{aligned} \tag{1}$$

In order to ensure low current consumption and a certain degree of noise sensitivity, the feedback current I_{FB} is required to be no less than $5\mu A$ [27]. Noise sensitivity and output voltage accuracy will increase with the increase of I_{FB}

TABLE 1: List of node peripherals.

Component	Model	Quantity	Description
STM32 development board	STM32L151C8T6	1	Main control chip, data processing, storage, and other operations
DC-DC converter	TLV62565	1	DC 5 V lithium battery power supply voltage is reduced to DC 3.3 V processor normal working voltage
Analog-to-digital converter (ADC)		1	Collect gas sensor data and battery power
Gas sensor	MQ-2	1	Analog output sensor for detecting the concentration of liquefied petroleum gas and alcohol (C3H6O)
Gas sensor	MQ-7	1	Analog output sensor for detecting carbon monoxide (CO) concentration
Temperature and humidity sensor	DHT11	1	Analog output sensor for detecting ambient temperature and humidity
LoRa communication module	SX1276	1	Wireless radio frequency module chip, upload the data in child nodes to the gateway
Antenna	SMA-KA	1	Increase LoRa RF frequency, dbm
LEDs		3	Alarm, normal working operation indication, low battery reminder, operation status indication
Buzzer		1	Fire alarm and low battery reminder

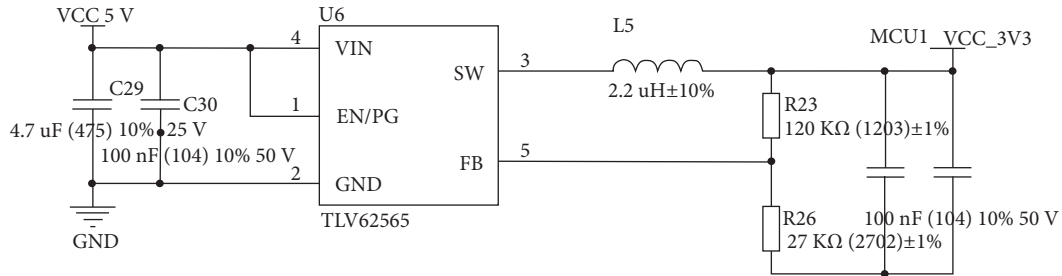


FIGURE 5: Schematic diagram of the step-down circuit.

and at the same time will increase the loss of output current. Due to the maximum duty cycle limitation, if input voltage is too low, output voltage will exceed the regulation range. To ensure that the circuit can work normally, V_{OUT} should be less than the product of minimum input voltage V_{IN_MIN} and the maximum duty cycle D_{MAX} , as shown in

$$V_{OUT} \leq V_{IN_MAX} \cdot D_{MAX}. \quad (2)$$

3.2. LoRa Gateway Hardware Implementation. The embedded gateway is composed of an industrial control core board and an industrial control backplane. The hardware structure diagram is shown in Figure 6. The industrial control core board is composed of i.MX6Ultralite embedded processor 256 MB Nand Flash and DDR3 [28]. This series of embedded processors uses a high-performance ARM Cortex-A7 32 bit core architecture with a frequency of up to 792 MHz [29]. The industrial control backplane includes LoRa gateway module and 4G LTE mobile data board, two 100M Ethernet ports, one CAN bus port, and RS485 bus port. Adopt 12 V DC power supply for power supply [30].

The embedded gateway is used to control data reception, processing, storage, and information forwarding. Collect the data uploaded by each node through the LoRa gateway module, check and analyze each data packet, and forward the correct data packet to the cloud server through the 4G module.

3.3. Software Implementation. The software program in this paper is mainly divided into three parts, as shown in Figure 7. The first is subnode program, which consists of sensor data acquisition, low power software optimization, and LoRa communication. The second is gateway program, which realizes the functions of receiving data from each child node, packaging, and uploading to servers. The third is server program, which realizes data query, real-time data display, and other functions. Besides, users can query real-time data by remote user monitoring system.

3.3.1. LoRa Node Software Design. The embedded software of child node is developed in Keil Embedded Development Tools integrated environment, and it is necessary to develop application layer programs and hardware drivers for peripheral devices. The procedure is divided into three parts. First, STM32 internal ADC collects sensor data into data. The working mode of carbon monoxide sensor, temperature and humidity sensor, and combustible gas sensor is to actively send data to single-chip microcomputer. In order to facilitate management, the single-chip microcomputer enables and disables ADC acquisition pins to control whether data is input. The software flow chart in MCU is shown in Figure 8.

After packaging the sensor data, the kernel and various peripherals are initialized by MCU. After the initialization is successful, the node sends a handshake signal to gateway,

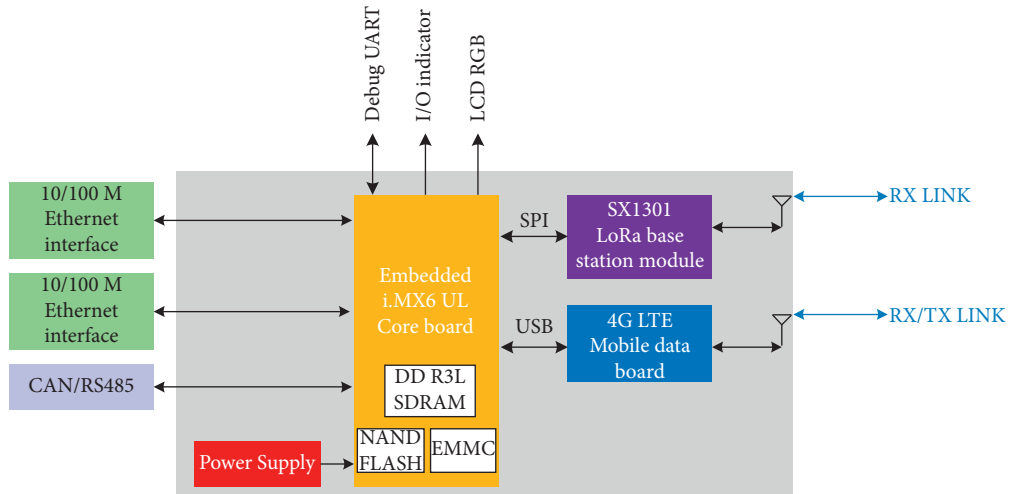


FIGURE 6: Gateway hardware structure diagram.

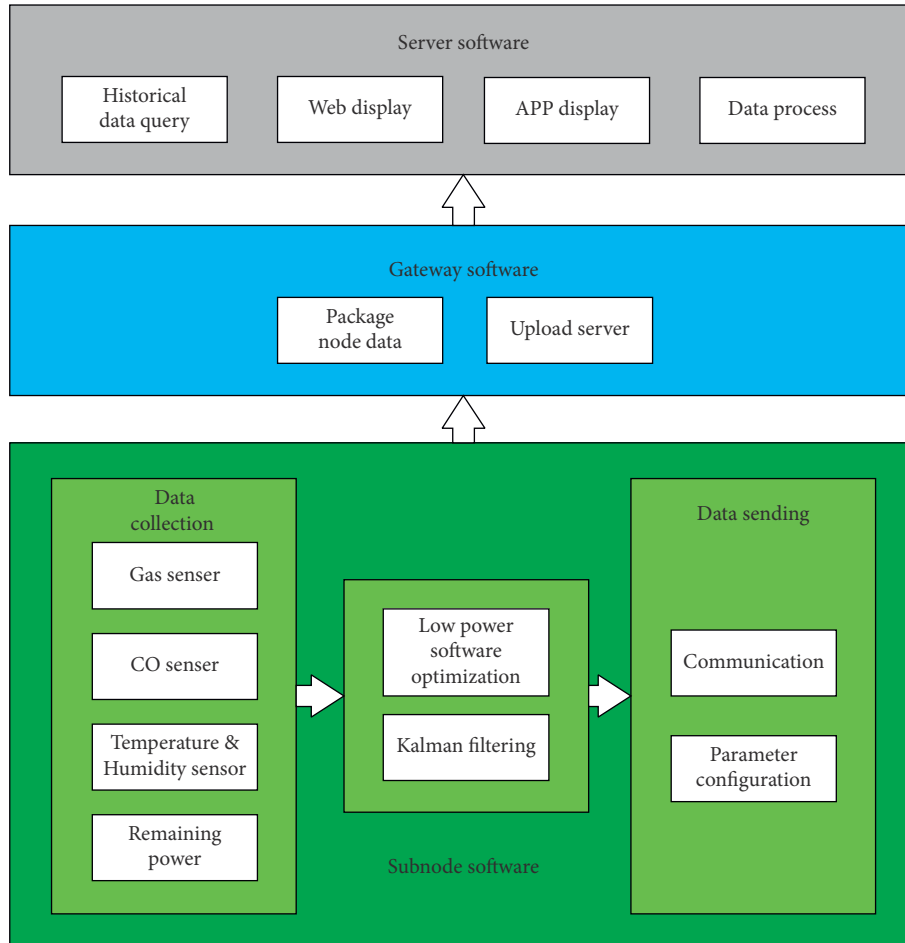


FIGURE 7: System software structure diagram.

asking the gateway whether to send configuration data, waiting for gateway to send data, or waiting for timeout to enter the low power consumption mode. If the gateway sends data, SX1278 clears interrupt and enters the receiving mode. In the receiving mode, MCU parses the message sent

by the gateway according to a certain protocol and judges the correctness for data frame format of messages. The protocol frame format is shown in Table 2. After receiving the data and enabling new parameters, enter STOP mode. If the wait times out, it directly enters STOP mode.

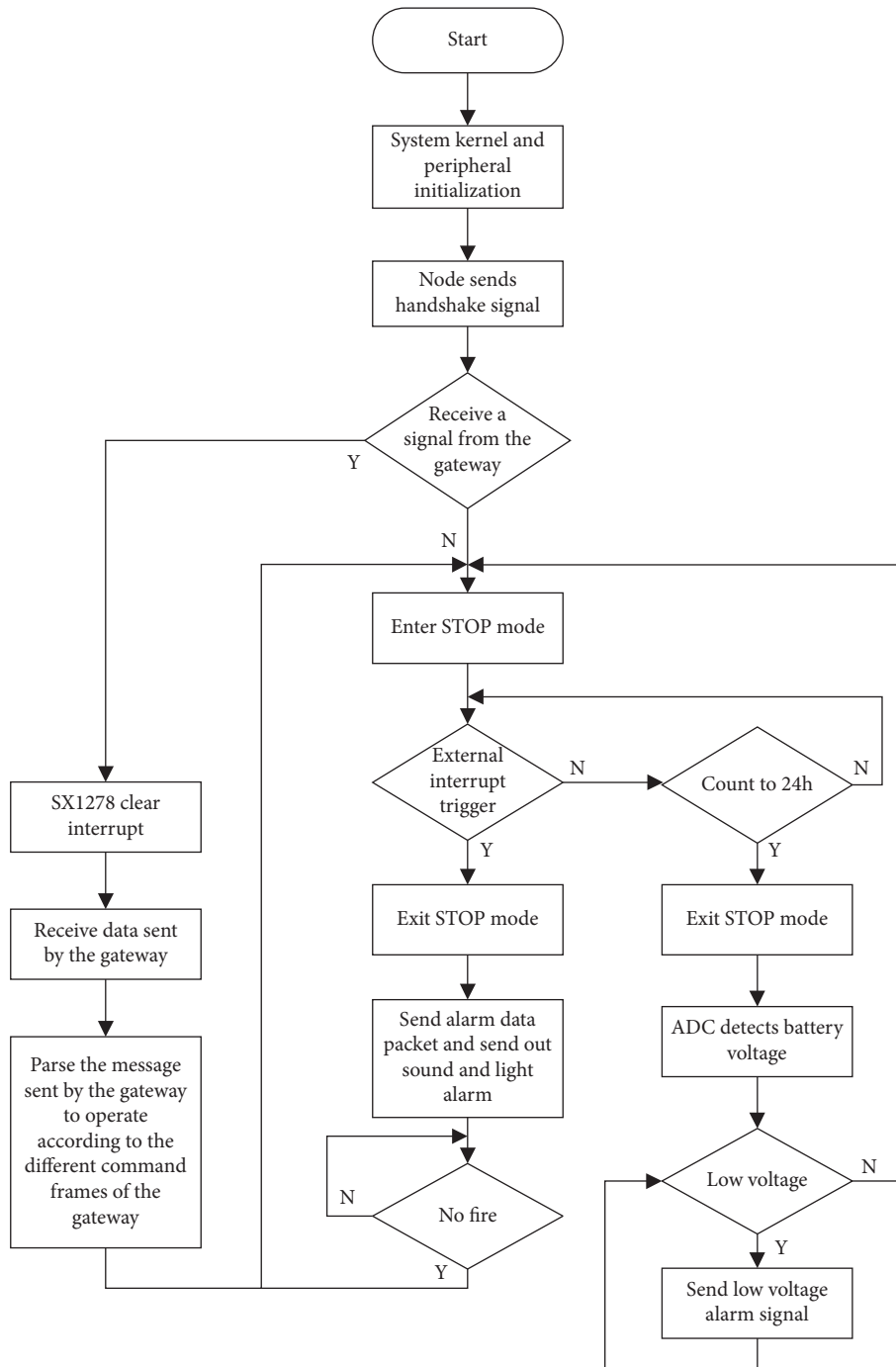


FIGURE 8: Node software flow chart.

TABLE 2: Protocol frame format table.

Frame content	Occupied bytes	Parameter range	Remark
Frame header	1		Corresponding character “@”
Target node ID	2	0x0000 – 0xffff	0xffff is a broadcast signal, all nodes respond
Configuration mode frame	1	0: flash 1: read	0 × 00: Set the wake-up time of nodes and store the parameters in flash 0 × 01: Immediately request the node to send current sensor data
Configuration data frame	2	0-65535	Analyze different data according to different configuration mode frames
Reserved byte	10		Reserved for future use
Frame tail	1		Corresponding character “\$”

In STOP mode, MCU detects the current smoke concentration by sensors. If the sensor detects a fire, it immediately sends a pulse signal to MCU interrupt IO port and uses an external interrupt to wake up MCU and set the SX1278 chip to send mode, continuously send alarm signals to gateway, and send out audible and visual alarm signals. If there is no fire, MCU maintains STOP mode state. The node wakes up every 24 h to detect the current system battery voltage. When the system detects that the current voltage is less than a certain value, it sends a low battery alarm signal to the gateway to remind staff to replace battery to ensure the normal operation of nodes.

3.3.2. Gateway Software Design. According to the flowchart shown in Figure 9, the gateway is in a passive receiving state in this system. When LoRa gateway receives the data packet of LoRa node, it parses the data packet. The analysis content includes CRC check bit, the node ID that sends data packet, the destination address of data packet, the data collected by sensor, and RSSI value of data packet.

After the gateway receives data packet sent from nodes, it stores data in an array with a length of 16 bytes. For the data packets stored in array, CRC check is performed first, if CRC check succeeds. Then analyze the data packet according to self-defined protocol, and discard the data packet if CRC check fails.

3.4. Application Layer Monitoring System. The application layer monitoring system mainly completes the real-time monitoring of measured data, alarm management, and viewing historical data. Equipment monitoring mainly monitors the status and data information of equipment as shown in Figure 10. The platform can monitor real-time data collected by sensors under each node and can display the historical data curve in current period of time. It can be searched according to provided conditions. Alarm information management is mainly to summarize and display the alarm information generated by system. The generation of alarm information is currently mainly realized by the trigger of system. When the trigger detects that real-time data of sensing point exceeds threshold set by administrator, it will start the trigger and send out an alarm message. The ID number, geographic location, and data name of alarm node are provided in alarm details, so that users can quickly locate the source of alarm. According to the state of alarm information, it is divided into three states: unprocessed, processed, and marked. The system administrator or user handles the alarm according to specific conditions of alarm. The historical data mainly records the historical data of equipment. The user can query historical data based on node ID number and sensor name under the node in combination with time.

4. Experimental Results and Analysis

Based on the abovementioned hardware and software platform, the system is tested for alarm response and monitoring. Before deploying system, the performance of

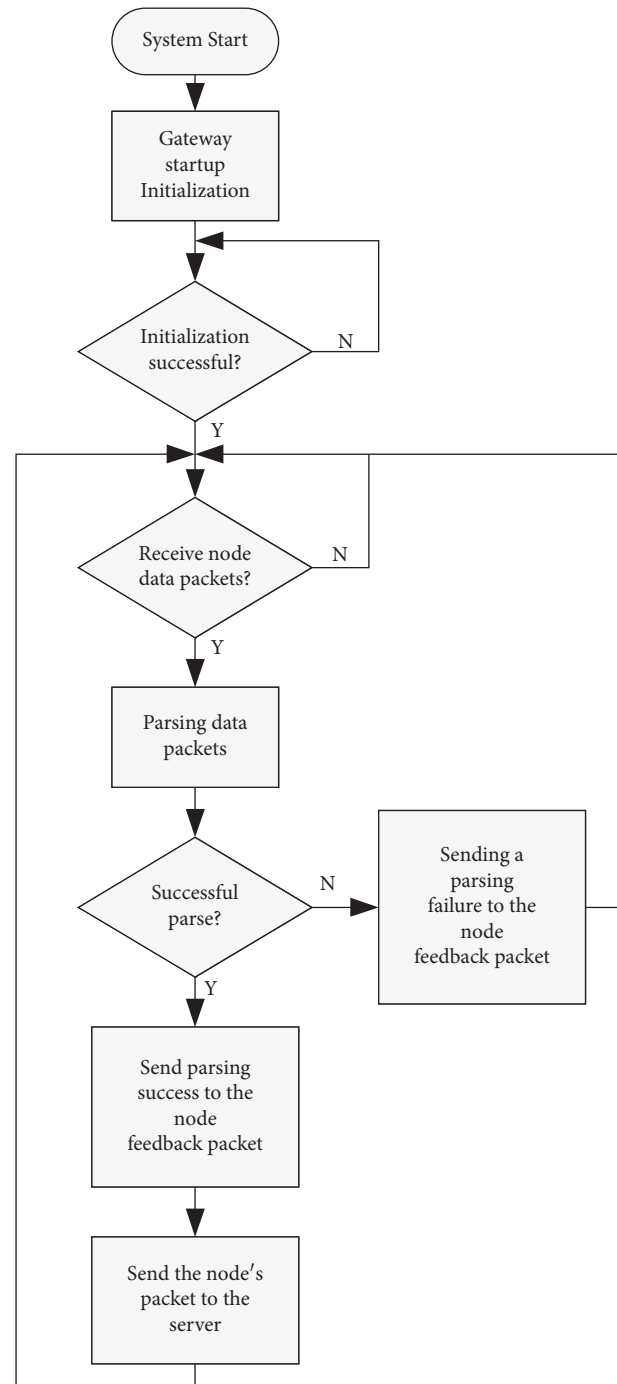


FIGURE 9: Gateway software logic flow chart.

system operation and different interference environments was tested. At the same time, the optimal transmission power of each node is selected to ensure that the node can stably receive data packets while operating with the lowest power consumption, thereby prolonging the service life of nodes.

4.1. Sensor Correction Experiment. When the concentration of carbon monoxide in air reaches 50 ppm, it is the maximum content allowed for an adult to be in it. When it

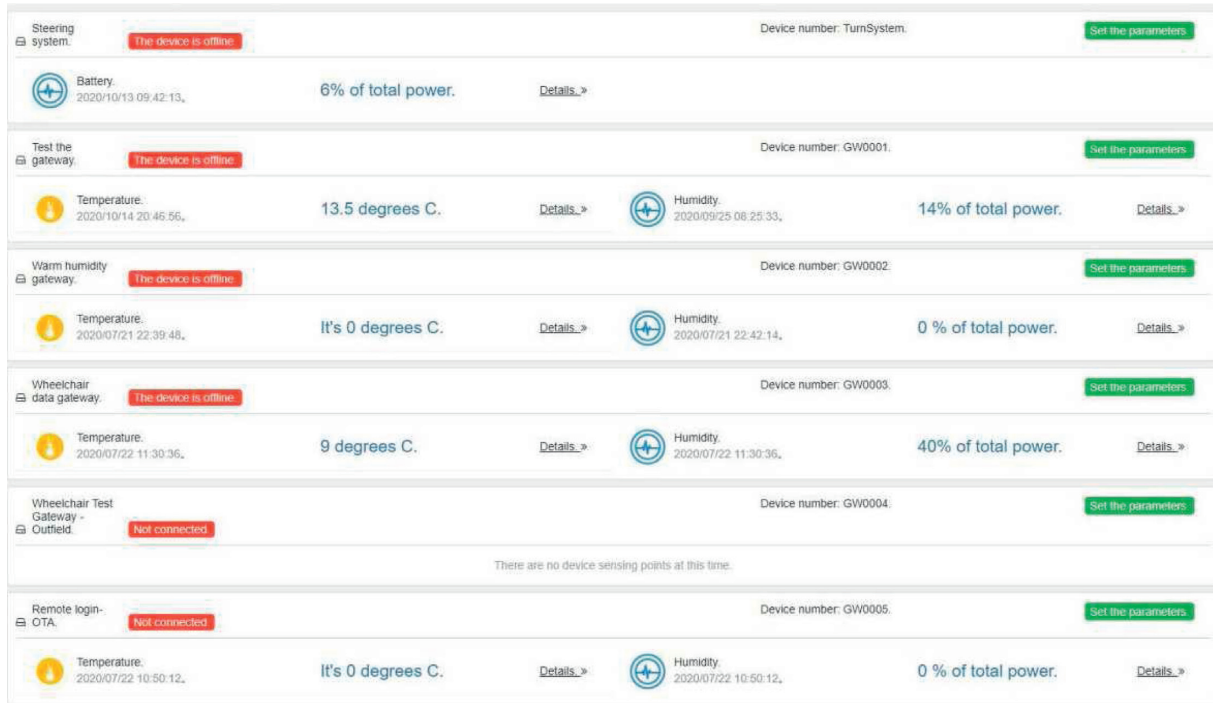


FIGURE 10: Device management interface of remote monitoring platform.

exceeds 200 ppm, the human body will have symptoms such as headaches and nausea. Therefore, when the carbon monoxide concentration is set to 50 ppm, it is alarm threshold. However, since Message Queue (MQ) series sensors use analog voltage output, it is necessary to refer to the sensitivity characteristic curve to convert analog voltage value into a value in ppm. Certain errors will occur during the conversion process. In order to improve the accuracy of value as much as possible, the carbon monoxide concentration was measured several times, and the experimental results are shown in Figure 11. The x-axis is analog voltage, and y-axis corresponds to ppm value of CO. It can be obtained by curve fitting that, in a ventilated room, the carbon monoxide concentration is 9.5 ppm and smoke concentration is 59.7 ppm.

4.2. Impact of Environment on the LoRa Communication Quality. The quality of communication environment will affect the quality of link transmission. When working in an environment full of irrelevant electromagnetic waves, the accuracy of receiving correct data packets will decrease accordingly. Thus, it is necessary to analyze the influence of indoor electromagnetic wave strength on communication quality. The communication quality of LoRa can be evaluated by received signal strength RSSI. Select two representative locations for testing in order to more realistically reflect the test results and study how different environments will affect the test results. The first experimental site is Numerical Control workshop of Nanjing Institute of Technology. This is a 123*38m processing workshop as shown in Figure 12(a)). Since electromagnetic harmonics are generated during the operation of machine tools, they are

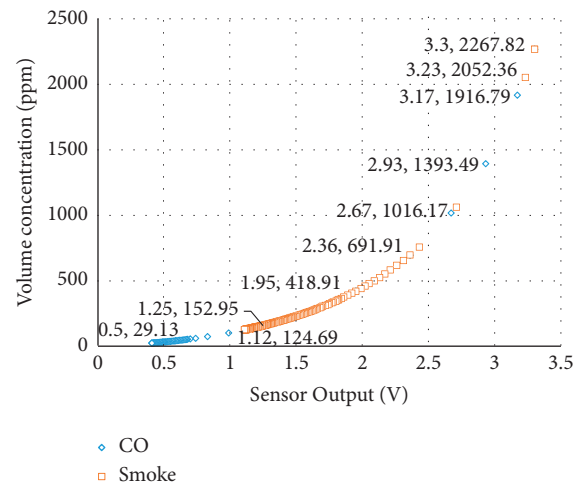


FIGURE 11: Volume concentration and analog voltage conversion.

the main source of communication interference, making it a very noisy experimental environment. The second experimental environment is school library, as shown in Figure 12(b)).

Set the communication parameters to SF=7, BW = 125 kHz, Freq_tx = 470.7 MHz, so that the transmit power is slowly increased from -4. When defining the interval dm between nodes and gateway, the minimum transmit power when the rate of packet reception of data packet received by the gateway within 10 minutes is higher than 95% is optimal transmit power. In order to ensure accuracy and authenticity of experiment, each test environment selects 8 locations for testing. Each test point is separated by a certain distance, and the optimal transmit



FIGURE 12: (a) Test environment 1. (b) Test environment 2.

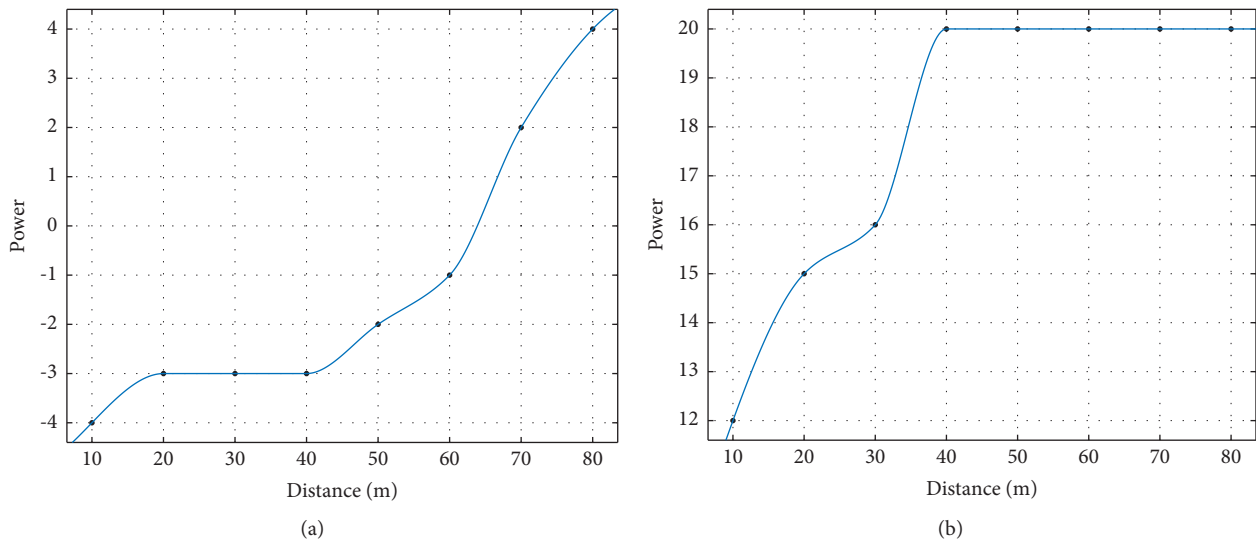


FIGURE 13: (a) Parameter configuration relationship of environment 1. (b) Parameter configuration relationship of environment 2.

power of each distance is measured. In both environments, the gateway is placed in the center of room, and the node is fixed on a wall at a certain distance from the ground. At the same time, if the gateway receives correct data packet, it measures its RSSI.

The experimental results of environment one and environment two are shown in Figures 13(a) and 13(b)). The best transmission power selection model is established by MATLAB curve fitting. It can be clearly seen from the figure that, in a noisy environment, it is necessary to greatly increase the transmission power to ensure normal channel communication. As shown in Figures 14(a) and 14(b)), the received signal strength is also significantly related to the excellence of environment. In a space with less interference, when the node is 40 m away from the gateway, only the transmit power needs to be set to -3, and the average received signal strength is maintained at about -104.12 dBm. In a space with large interference, when the node is at the same distance from the gateway, the transmit power needs to be set to 20 to ensure that the packet reception rate is higher than 95%, and the average received signal strength is only about -57.5 dBm.

5. Distributed Fire Monitoring System Test

8 LoRa nodes and 1 gateway are deployed in a star network, and the node transmitting power is set to the optimal transmitting power of corresponding distance. Perform comprehensive tests of sensor collection data, node-gateway communication, gateway-server communication, and remote user monitoring system display. Through human intervention, the release of carbon monoxide and smoke and increasing the temperature simulate occurrence of a fire, further testing the feasibility of the system. When the index is not detected to exceed standard, the remote user monitoring system will not issue an alarm. In the event of a fire, node immediately sends out an audible and visual alarm signal, and the remote user monitoring system sends out an alarm message at the same time. In order to obtain the accuracy of experimental data, 8 nodes were tested.

It can be seen from Table 3 that after sensor verification and selection of appropriate transmitting power, under normal conditions, the temperature measured at node 8 is 19.4°C, the carbon monoxide content is 11.6 ppm, and the

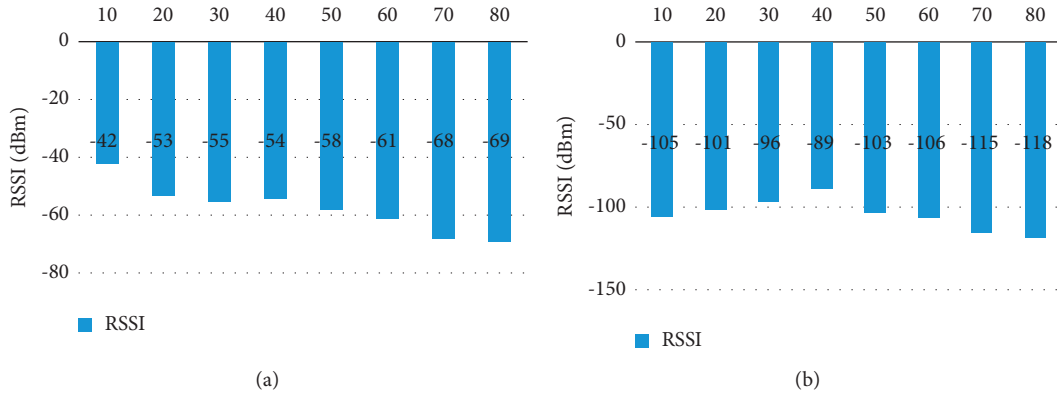


FIGURE 14: (a) RSSI at different distances in environment 1. (b) RSSI at different distances in environment 2.

TABLE 3: Alarm accuracy.

Sampling node	Temperature/(°C)	Carbon monoxide/(ppm)	Smoke/(ppm)	Send alarm
1	58.2	316.8	128.4	Y
2	55.6	20.2	127.4	Y
3	54.8	19.6	387.2	Y
4	48.3	24.9	244.6	Y
5	42.8	156.9	259.1	Y
6	35.2	53.0	257.8	Y
7	26.6	25.1	154.9	N
8	19.4	11.6	102.4	N

smoke is 102.4 ppm. Then randomly increase the temperature, carbon monoxide, and smoke concentration values, and record the alarm information. Node 7 first improves the carbon monoxide. Judging from the data displayed by the monitoring platform, all three values have increased. After data fitting calculations, an alarm will be given. It was observed from the monitoring platform that the accuracy of alarm reached 100%, and the system was operating stably. This realizes the collection of real-time data, the distributed wireless communication function, and the alarm function after threshold is exceeded, meeting the needs of fire alarm.

In summary, the quality of LoRa wireless communication is greatly affected by the environment, and the transmission power should be reasonably increased in a large electromagnetic interference environment. In practical applications, in order to prolong the service life of the node as much as possible, it was found in the test that when the node is close to the gateway, too high transmission power will cause data packet reception errors, so you should choose the appropriate one according to different transmission distances. The transmission power can also ensure good alarm accuracy.

6. Conclusion

This paper proposes an indoor distributed fire alarm system based on LoRa for IoT. The system realizes the deployment of distributed nodes and gateways in a star network structure by LoRa communication technology. The node collects

environmental data and synchronizes data to remote monitoring system in real time. In the event of a fire, the node monitoring system can issue an alarm in real time. The system has the characteristics of easy deployment and long service life. The sensor filtering and weighted fusion algorithm can accurately determine whether a fire has occurred, and the remote user monitoring system can monitor the system status of each node in real time. At the same time, the node implements regular data transmission and low power sleep and energy-saving modes to facilitate the long-term use of devices.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. SJCX21_0921), School-level Scientific Research Fund of Nanjing Institute of Technology (No. QKJ201808), and 2020 Jiangsu Province Industry-University-Research Cooperation Project (No. BY2020040).

References

- [1] W. Yunlong, "Current status and improvement of fire protection supervision of high-rise buildings in my country," *Green Building Materials*, vol. 309, no. 9, pp. 165–166, 2021.
- [2] A. Tzounis, N. Katsoulas, T. Bartzanas, and C. Kittas, "Internet of Things in agriculture, recent advances and future challenges," *Biosystems Engineering*, vol. 164, pp. 31–48, 2017.
- [3] X. J. Xing, J. C. Song, L. Y. Lin, M. Q. Tian, and Z. P. Lei, "Development of intelligent information monitoring system in greenhouse based on wireless sensor network," in *Proceedings of the 2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, pp. 970–974, IEEE, Changsha, China, July 2017.
- [4] L. Yunhong and Q. Meini, "The design of building fire monitoring system based on zigbee-wifi networks," in *Proceedings of the 2016 Eighth International Conference on Measuring Technology and Mechatronics Automation*, pp. 733–735, IEEE, Macau, China, March 2016.
- [5] X. P. Shen, X. Wang, and M. Jia, "Design and implementation of traffic information detection equipment based on Bluetooth communication," in *Proc. IEEE Inf. Technol. Netw. Elect. Autom. Control Conf.*, pp. 1595–1601, IEEE, Chengdu, China, December 2017.
- [6] S. Basu, S. Pramanik, S. Dey, G. Panigrahi, and D. K. Jana, "Technology, Fire Monitoring in Coal Mines Using Wireless Underground Sensor Network and Interval Type-2 Fuzzy Logic Controller," *International Journal of Coal Science & Technology*, vol. 6, no. 2, pp. 274–285, 2019.
- [7] X. Zhang, J. Du, C. Fan, D. Liu, J. Fang, and L. Wang, "A wireless sensor monitoring node based on automatic tracking solar-powered panel for paddy field environment," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1304–1311, 2017.
- [8] M. Iqbal, A. Y. M. Abdullah, and F. Shabnam, "An application based comparative study of LPWAN technologies for IoT environment," in *Proceedings of 2020 IEEE Region 10 Symposium (TENSYP)*, pp. 1857–1860, Dhaka, Bangladesh, 2020.
- [9] G. Roque and V. S. Padilla, "LPWAN based IoT surveillance system for outdoor fire detection," *IEEE Access*, vol. 8, pp. 114900–114909, 2020.
- [10] D. Patel and M. Won, "Experimental study on low power wide area networks (LPWAN) for mobile Internet of Things," in *Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, Sydney, NSW, Australia, June 2017.
- [11] T. Defeng, L. Shixing, X. Wujun, and Z. Yongming, "A fire monitoring system in ZigBee wireless network," in *Proceedings of the 2010 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*, pp. 48–51, IEEE, Huangshan, China, October 2010.
- [12] R. B. Sørensen, D. M. Kim, J. J. Nielsen, and P. Popovski, "Analysis of latency and MAC-Layer performance for class A LoRaWAN," *IEEE Wireless Commun. Lett.*, Mar. vol. 6, no. 5, pp. 566–569, 2018.
- [13] S. Barrachina-Muñoz, B. Bellalta, T. Adame, and A. Bel, "Multi-hop communication in the uplink for LPWANs," *Computer Networks*, vol. 123, pp. 153–168, 2017.
- [14] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 60–67, 2016.
- [15] H. Chu, Z. Xie, and H. Jiang, "Hybrid intelligent monitoring network based on Ad hoc and wireless sensor networks," *Information Technology Journal*, vol. 10, no. 4, pp. 849–855, 2011.
- [16] L. Feltrin, C. Buratti, E. Vinciarelli, R. De Bonis, and R. Verdone, "LoRaWAN: evaluation of link- and system-level performance," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2249–2258, 2018.
- [17] J.-G. Kang, D.-W. Lim, and J.-W. Jung, "Energy-Efficient forest fire prediction model based on two-stage adaptive duty-cycled hybrid X-MAC protocol," *Sensors*, vol. 18, no. 9, p. 2960, 2018.
- [18] H.-C. Lee and K.-H. Ke, "Monitoring of large-area IoT sensors using a LoRa wireless mesh network system: design and evaluation," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 9, pp. 2177–2187, 2018.
- [19] Y. Ma, L. Zhao, R. Yang et al., "Development and application of an atmospheric pollutant monitoring system based on LoRa-Part I: design and reliability tests," *Sensors*, vol. 18, no. 11, p. 3891, 2018.
- [20] O. Perešini and T. Krajčovič, "More efficient IoT communication through LoRa network with LoRa@ FIIT and STIOT protocols," in *Proceedings of the 2017 IEEE 11th International Conference on Application of Information and Communication Technologies (AICT)*, pp. 1–6, IEEE, Moscow, Russia, September 2017.
- [21] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: an overview," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 855–873, 2017.
- [22] Y. Song, J. Lin, M. Tang, and S. Dong, "An Internet of energy things based on wireless LPWAN," *Engineering*, vol. 3, no. 4, pp. 460–466, 2017.
- [23] J. Chen, "Narrowband Internet of things: implementations and applications," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2309–2314, 2017.
- [24] H. Ruotsalainen, J. Zhang, and S. Grebeniuk, "Experimental investigation on wireless key generation for low-power wide-area networks," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1745–1755, March 2020.
- [25] M. Rizzi, P. Ferrari, A. Flammini, and E. Sisinni, "Evaluation of the IoT LoRaWAN solution for distributed measurement applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3340–3349, 2017.
- [26] B. Sarwar, I. S. Bajwa, N. Jamil, S. Ramzan, and N. J. S. Sarwar, "An intelligent fire warning application using IoT and an adaptive neuro-fuzzy inference system," *Sensors*, vol. 19, no. 14, p. 3150, 2019.
- [27] E. Tanir, K. Felsenstein, M. Yalcinkaya, and E. S. Science, "Using Bayesian methods for the parameter estimation of deformation monitoring networks," *Natural Hazards and Earth System Sciences*, vol. 8, no. 2, pp. 335–347, 2008.
- [28] L. Tuo, X. Zheng, X. Chen, and P. Research, "A new online exhaust gas monitoring system in hydrochloric acid regeneration of cold rolling mills," *Environmental Science and Pollution Research*, vol. 24, no. 35, pp. 26919–26926, 2017.
- [29] T. Voigt, M. Bor, U. Roedig, and J. Alonso, "Mitigating Inter-network Interference in LoRa Networks," in *Proceedings of the 2017 International Conference on Embedded Wireless Systems and Networks*, Uppsala, Sweden, February 2017.
- [30] X. Zhang, M. Zhang, F. Meng, Y. Qiao, S. Xu, and S. Hour, "A low-power wide-area network information monitoring system by combining NB-IoT and LoRa," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 590–598, 2019.

Review Article

Futuristic Cyber-Twin Architecture for 6G Technology to Support Internet of Everything

Sapna Juneja ¹, Mamta Gahlan ², Gaurav Dhiman ³ and Sandeep Kautish ⁴

¹IMS Engineering College, Ghaziabad, Uttar Pradesh, India

²Maharaja Surajmal Institute of Technology, Delhi, India

³Department of Computer Science, Government Bikram College of Commerce, Patiala 147001, Punjab, India

⁴LBEF Campus, Kathmandu, Nepal

Correspondence should be addressed to Sandeep Kautish; dr.skautish@gmail.com

Received 4 September 2021; Accepted 27 September 2021; Published 6 October 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Sapna Juneja et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid growth of Internet of Everything, there is a huge rise in the transportable Internet traffic due to which its associated resources have exceptional obstacles which include reliability, security, expandability, security, and portability which the current available network architectures are unable to deal with. In this paper, an IoT centric cyber-physical twin architecture has been proposed for 6G Technology. The cyber-twin technology helps out in serving stronger communication and also contains several features that help out in assisting communication like maintaining a log record of network data and managing all digital assets like images, audio, video, and so forth. These features of the cyber-twin technology enable the proposed network to deal with those exceptional obstacles and make the system more reliable, safe, workable, and adaptable.

1. Introduction

The population of the world is growing like a bomb and the people are now connected with each other with the help of millions of devices. The available networks will be insufficient to accomplish the rising demand of billions of devices thereafter. Thus, the presumed future network should be able to gratify the intense network traffic and its associated services. In the current times, Internet of Everything seems to be a future network system that can attain rational connections between humans and devices with the help of various technologies like Machine Learning and 5G. An Internet of Everything enabled network architecture to reinforce the complex features of data like collecting pervasive data, data clustering, refining, and lastly allocating it for utilization.

2. Literature Review

The current network architecture suffers from the scalability problem, that is, not being able to fulfill the tremendously increasing demand of Internet services and devices. Further,

in order to secure the network, all the accessed devices can be kept under observation to detect any unauthorized access but it is not feasible to keep a watch on every user accessing the network. Thus, these problems lead to the security breach in the network. Another major problem faced by current available networks is that maintaining collaboration and synchronization among various resources controlled by multiple network service providers is very burdensome which finally affects the quality of the service of the network. All these problems faced by available network architectures are unfavorable and thus hinder the growth and development of the services provided by the available networks. For dealing with such issues and challenges, many researchers tried to propose various kinds of network architectures, each with its own pros and cons. Few of them are named as Named Data Networking, DONA, Content Centric Networking, and so forth but still these networks are not self-sufficient to support huge traffic as these networks are incompatible with current IP networks [1].

One of the architectures named MobilityFirst [2] was developed to deal with current issue related to network but it

was lacking the security of the whole network. Another architecture proposed was Expressive Internet Architecture [3] which was better than the former in terms of network security as it was incorporating various protocols used to control and manage the network but it also has some pitfalls associated with it like difficulty in network expandability and relatively high cost of establishing and managing the network. ChoiceNet [4] was developed to facilitate choices and the financial relationships among the various entities of the network but it is not much efficient for complex real-time services. Further in the queue, a cloud-based architecture had also emerged with the potential of sharing network resources wisely and managing the immense data gathered from IoT or IoE devices such as CloNe, Cloud Integrated Networks [4], and Nebula. Nevertheless, these networks also face performance issues due to the low computational ability of the connected devices. Another problem associated with these networks is the lack of mobility as it is incapable of serving the mobile devices, although the mobile devices are much trending nowadays. One of the types of cloud-based network architectures is CloNe architecture. It was beneficial for cloud customers as it helps to connect network and data domains. Apart from that, network resources are spread properly in the whole network among various devices, thus providing better connectivity to the end user. These networks are unable to handle mobility in the network though. The Cloud Integrated Network architecture [5] fixes the cloud of the network edge to the network itself, thus providing the optimal throughput. It is efficient in the terms that it manages the resource distribution and utilization credibly but it is not able to deal with mobility issue of the devices in the cloud. Hence, a variety of network architectures have been defined by the researchers but still they are not appropriate for implementing as they encounter multiple challenges like mobility, expandability, security and so forth. In this research paper, cloud-based cyber-twin architecture [6] has been proposed as the futuristic network. By using cyber-twin architecture, we can implement a novel architecture of network that must be able to deal with the issues of mobility, expandability, security, and reliability. Table 1 represents the work done by various researchers in the field of cyber-twin technology:

3. The Concept of Cyber-Twin

Because of the sudden rises in the digitalization of the whole world in Industry 4.0, the concept of cyber-twin arose. The main motive behind the development of cyber-twin is to support Industry 4.0 in such a manner that the production environment becomes more flexible, productive, and resource-efficient. In order to make the production more fast and progressive, the most important constraint is the minimization of the response time to the customer's query and this can be fulfilled only when there exists a network that is fast enough to manage everything within a blink [15]. This fast network can be supported using cyber-twin. Before moving to the concept of cyber-twin, some basic terminologies need to be studied for understanding the purpose of cyber-twin.

3.1. Cyber-Physical System. Cyber-physical systems are the systems that act as a base for Industry 4.0. The structure of the cyber-physical system is a collection of hardware and software that work together with embedded system. Embedded systems [16] are being used to manage the sensors and their states present in the cyber-physical systems. Cyber-physical systems can be summed up as combination of hardware and software, where the software affects the behavior of the system. Further, the cyber-physical system can visualize the physical components in real-time environment and hence can be used for simulation of the physical components. Any function is being performed by the cyber-physical system by accessing the recorded data.

3.2. Digital Twin Paradigm. A digital twin is a virtualization of any physical system that keeps on updating every time by analyzing the performance of the previous one. If the original physical system and its twin copy are integrated with each other and share the data and information, the twin is known as digital twin [17]. So, in such scenario, if the changes are being done in one system, these changes will automatically reflect to the other system and vice versa. Ontology based digital twins can be used to implement and manage the digital twins in current data ecosystem. Figure 1 shows the structure of digital twin.

3.3. Cyber-Physical Twin. Cyber-physical twin system is a key methodology that entitles the products which are viable and reliable in terms of their services. The system is composed of two-level hierarchies, cyber-twin and physical twin. Both twins associate with each other for completing the life cycle of cyber-physical system. The physical twin is made up of actual modules, that is, sensors and actuators. Sensors are being used by the physical twin to instigate the impact of environment [18]. The data thus collected from the sensors has been sent to the cyber-twin which further analyzes the data and, depending on this analysis, provides feedback to the physical twin and the user. Communication between the user and the twins can be achieved with the help of an interface. The interface is also used for controlling the behavior of the physical twin and analyzing the reaction of the cyber-twin. The next section contains the behavior of cyber-twin and physical twin and also discusses the mode of interaction between them [19].

3.3.1. Cyber-Twin. The cyber-twin is pragmatic approach of any physical component of the cyber-physical twin that incorporates the response of the physical component. Hence, the cyber-twin behaves much similar to the physical component. The core of the cyber-twin embraces a model that contains all the relevant information and its associated relationships. The data in this model has been stored by the available sensors of the system and can be used later on for performing computations. Thus, this model of cyber-twin contains a database that accommodates the data used for carrying out the operations by physical twin [20]. The cyber-twin generates an algorithm which transforms the behavior of the physical components.

Further, by applying the simulation technique, the present situation and the upcoming situation and the functions of the physical twin can be predicted.

3.3.2. Physical Twin. A physical twin is a term similar to digital twin, which elaborates the physical systems or products of the real world, which further is replicated by digital twin. The physical twin is enabled with the sensors that are provided to keep track of the functions and the behavior of the components of the product during the usage phase of the product. Further, these physical components work in real-time environment with the help of actuators. An interface is also added to the physical twin to perceive the interactivity between the physical twin and cyber-twin.

3.3.3. Communication between Physical Twin and Cyber-Twin. The communication between the physical twin and cyber-twin must be made in such a way that there should be no break in the communication between them. Thus, real-time communication should be incorporated among the twins. In order to pursue this real-time communication between the twins, various protocols like MQTT and so forth can be implemented. The steps followed for the communication between the physical twin, cyber-twin, and the end user can be elaborated as follows:

- (1) The requirements must be gathered as input that must meet the choice of individual user. These requirements further help out in the configuration of cyber-physical twin system.
- (2) The next step is to configure the cyber-physical twins precisely and carefully. This configuration is a two-step process in which the first step is the implementation of the sensors, which further forms the physical part of the cyber-physical twin. In the second step, the virtual twin of the respective cyber-twin is created.
- (3) Both processes are connected together and reconfigure the database of the cyber-twin. This reconfiguration of the database is aligned by the sensors and the simulation models.
- (4) Further, sensors are being used to quantify the real physical influences of the environment.

4. Structure of Cyber-Twin Network

The proposed futuristic structure of cyber-twin network for 6G is composed of four parts: parent cloud, child cloud, user interface, and cyber-twin. Figure 2 shows the proposed structure of cyber-twin network.

The functions of each of the part of cyber-twin architecture are as follows:

- (1) Parent cloud: Various clouds are connected to each other to form the parent cloud. This parent cloud establishes the base of the network. Connection between these clouds can be of any type like via satellite or fiber optics. The difference between the

existing cloud and future cloud would be that they would not only be used for storing and computation of data; instead these parent clouds will also be used for calculation, caching, and sharing resources efficiently among various end users.

- (2) Child cloud: This structure is a three-level hierarchy where the topmost level is for parent cloud and the bottom level is for user interface. The child cloud lies in between the parent cloud and child cloud. Thus, child cloud is used to serve the functionality to the user through user interface as provided by parent cloud. Other than this, the parent cloud provides more features to the user in comparison to the child cloud. It can easily process out the user's request as it is adjacent to the user interface.
- (3) User interface: User interface here represents the interface that has been used to provide the connectivity between the clouds and users and further user refers to various devices connected to the network being used by the individual user. So, if any user wants to access the data or share the resources with any other user in the network, a request is made by user, which will be further sent to the child cloud and then to parent cloud. The user needed to be connected to the cloud all the times through cyber-twin for performing any computation, instead of being connected to the server.
- (4) Cyber-twin—It is the most important part of the futuristic network. It has been positioned adjacent to the child cloud and the communication it provides is different from the communication provided by the existing networks; thus it is capable of supporting many services which have not been provided by the present communication models [21].

5. Precedence of Cyber-Twin System over Peer-to-Peer Communication System

Futuristic cyber-twin network architecture supercedes the present peer-to-peer architecture in so many ways and the most important functionalities provided by the cyber-twin which the peer-to-peer system is unable to yield are the following:

- (a) Cyber-twin has the ability to keep a log record of the data of all the users connected via cyber-twin network which remains shielded as no external user is allowed to access the log record of cyber-twin, while, in the case of peer-to-peer architecture, various ISPs keep track of all the log records of the end user which further can be easily hacked by a hacker [15].
- (b) Cyber-twin is a way of representing any entity digitally. This entity can be either human being or any object. In the physical world, cyber-twin has been used to assist the entity. While assisting the user, the cyber-twin architecture requires the end user to get connected to the cyber-twin which further communicates with the cloud (child and parent) for

TABLE 1: A critical review of significant similar contributions from researchers.

S. no.	Author and year	Methodology	Findings
1.	[7]	Cloud operated cyber-twin architecture has been defined.	Architecture is very efficient for high speed networks but lacks privacy and security of the data.
2.	[8]	Digital twin architecture has been defined by taking the concept of data analytics and virtual reality.	Future industrial applications of digital twin architecture are explained. Various objectives that can be used for improvisation of digital twin technology are elaborated.
3.	[9]	A five-layered architecture of cyber-physical system is proposed which can further be implemented in Industry 4.0 for efficient production management.	The paper provides practical guidelines to manufacturing organizations for the implementation of cyber-physical systems for intelligent manufacturing process.
4.	[10]	Cyber-physical twins and their associated concepts are reviewed in the paper.	Advantages of cyber-physical twins are evaluated. The behavior and the usage of cyber-twins in real time environment are stated in the paper.
5.	[11]	In this paper, a four-layer reference model for digital twin has been proposed. This reference model contains guidelines for risk prevention and prediction using digital twin for enhancing the safety of employees at their workplace.	The proposed digital twin model is very useful for the organizations where employees live in high risk manufacturing environment as the proposed model is capable of developing virtual modeling environment for the manufacturing process, thus identifying the risks involved.
6.	[12]	Integration of discrete event simulation and digital twin has been explained in the paper and the challenges of contemporary logistics have been picked out.	Digital twins can be used for enhancing the functionality of DES for real-time simulation environment.
7.	[13]	An approach for designing digital twin technology for communication system has been used. Implementation of the model is done on Amazon Web Services.	The proposed system has the ability to act as the replica of the future power system.
8.	[14]	A digital driven framework for designing of products has been represented and a case study has been used to show the advantages of this proposed framework.	The proposed approach seems useful for redesigning of an existing product instead of designing the whole system from the beginning.

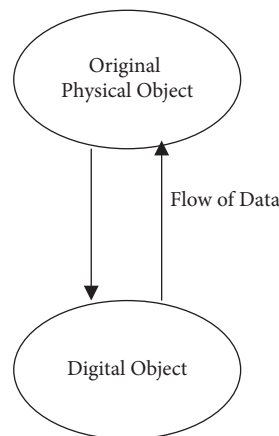


FIGURE 1: Structure of digital twin.

accomplishing the request done by the end user, while in the peer-to-peer network the end user directly connects to the cloud for obtaining the services [8]. Thus, in the former case, the user achieves better quality of service due to the cyber-twin's ability of knowing the user requirements well.

6. Advantages of Cyber-Twin System

If required resources of the cyber-twin networks will be available in proximity and if used efficiently, cyber-twin networks can have a revolutionary impact in the digital world.

The following section presents the benefits offered by the cyber-physical twin network to the entire community:

- (1) As the cyber-twin system gathers a lot of data for all the communication over the entire network, this collected data can further be utilized to enhance the properties of next system which is to be developed for the ease of use by the user and thus helps in developing a more better system that must be able to fulfill the user's requirements in all aspects [22].
- (2) In cyber-twin system, the lifecycle maintains a parallel two-phase process in which the first phase involves execution of the current cyber-twin system

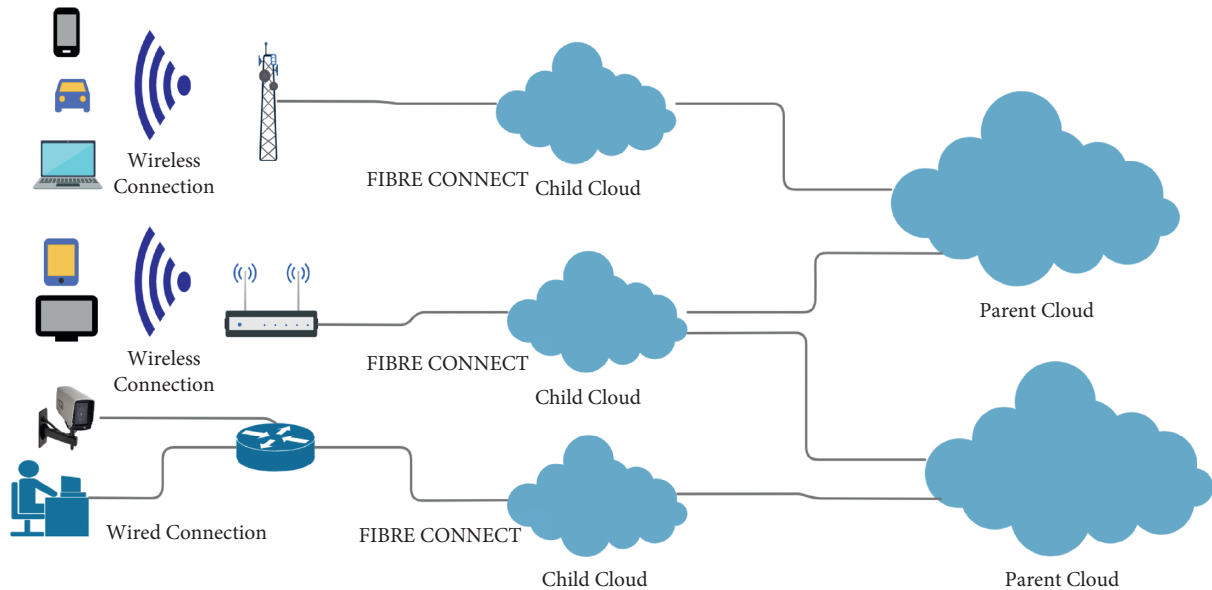


FIGURE 2: Structure of cyber-twin network.

and the second phase involves modifications in the existing system depending upon the current behavior of the system. Hence, two processes keep on continuing all the times which does not affect the throughput of the system.

- (3) In any of the traditional networks, all the hardware involved has some lifetime and once the lifetime gets over, the hardware should be replaced by the new component, which ultimately affects the performance of the system as, during the replacement process of the component, the whole communication process gets stopped. Meanwhile, in the case of cyber-twin system, the lifetime of any of the components of the whole network can be easily predicted and thus it can be replaced before reaching the end. As the cyber-twin system leads to more accuracy and better performance and thus prevents the system from getting failed due to component failure, it has more reliability in comparison to traditional networks.
- (4) As the number of mobile devices in the network is growing very fast, some mechanism is needed to take care of mobile data as the conventional network architecture is insufficient to serve mobile data because the present networks manage the data traffic by taking host address into account, which remains stationary. In the cyber-twin network architecture, every entity, either human or device, is represented by a digital ID not by host address which is mobile and thus cyber-twin network allows mobility of traffic for the convenience of users [23].
- (5) The cyber-twin network architecture provides more security to the data as it has several authentication processes running in the background which first confirms whether the user or the device is authentic and only then it allows the entity to access the data of

the cyber-twin network. Different security mechanisms are imposed on different entities like if a human being wants to access the data; the retina of the eye is scanned first to give access to the user which is somehow better than the security mechanism provided by current network architectures as these architectures do not contain such tough constraints on the access of the data.

7. Cyber-Twin Applications in Real-World Environment

This section of the review focuses on the application areas of cyber-twin in real-world environment ranging from healthcare and manufacturing to development of smart cities.

- (1) Healthcare: The most important application area of cyber-twin is the healthcare sector. As the healthcare sector is now growing exponentially due to the usage of IoT devices, the implementation of cyber-twin makes it possible to share the large amount of data collected from IoT devices at a very high speed which the traditional networks are not able to do with the same accuracy and speed. Secondly, cyber-twin can also be used for simulating the drugs' effects on the human body as well as planning and execution of medical therapeutics. Cyber-twin is also beneficial in AI-enabled healthcare for decision- and prediction-making especially in the situations where it seems difficult for the physicians to take the decision about any medical treatment for any patient. Digital twin has the ability to make the decisions based on real-life situations [24].
- (2) Manufacturing: Nowadays, manufacturing industry is looking for the system that can provide high connectivity and should be able to keep a track of the

products. These applications to manufacturing industry are being supported by cyber-twin. Digital twin has the ability to track the performance of the machines in manufacturing industry in real-time environment and also to predict the lifetime and future performance of the machine which is very important for performance and reliability measuring for the product development. Products can be easily tested by setting up a simulation environment by the cyber-twin [21].

- (3) Development of smart cities: The smart cities are now growing very fast because of the increase in IoT devices; people are coming more closer to each other for sharing information and, in this scenario, digital twins are required to connect all these communities together. The cyber-twin can help out in planning of smart cities by using AI techniques and can also be used to model smart buildings, smart traffic management, smart farming and smart livestock, and so forth [25].

8. Challenges of Cyber-Twin Technology

It is now apparent that cyber-twin requires various technologies to run in parallel with it and thus faces a lot of challenges. So, in order to deal with these challenges, the first task is to identify them. In this section, some challenges faced by digital twin technology will be discussed so that they must be taken into account in the development of the cyber-twin system.

- (1) The cyber-twin architecture requires infrastructure that must be compatible with the various upcoming technologies like AI and IoT that will run in collaboration with the cyber-twin for the effective running of the entire network. Without a compatible and well-defined infrastructure, the cyber-twin will fail in achieving its goal, resulting in a malfunction in the system.
- (2) The second challenge that appears is the data required for cyber-twin. The required data should be of high quality without any interruption in between. Interrupted and low-quality data leads to the degradation in performance of cyber-twin. Some analysis of the collected data should be there to check whether the data is perfect in all aspects and does not carry any uncertainty in it for the smooth functioning of cyber-twin architecture [26].
- (3) Another major challenge concerned with the cyber-twin architecture is the privacy of the sensitive data which it obtains from the multiple IoT devices connected to it. So, for dealing with the data security related to cyber-twin, some authentication and security mechanism should be applied at the IoT device to prevent it from unauthorized access, thus protecting the cyber-twin data [27].
- (4) Despite all these challenges, the next challenge that comes into light is that no standard design has been

available for the development of cyber-twin and if the system will be developed lacking standard design, there will be no uniformity throughout which will mislead the information flow, thus resulting in the imperfect system not capable of achieving its goal (A. [28]).

- (5) The last challenge faced by digital twin technology is that digital twin requires more maintenance and people are neither well equipped nor skilled in the maintenance of cyber-twin because of lack of availability of resources required for cyber-twin.

9. Conclusion and Future Work

The paper deals with the definition and the concepts of cyber-twin and the representation of structure of futuristic cyber-twin network. Further, the comparison between the cyber-twin network and peer-to-peer network has been elaborated to show how cyber-twin networks are better than peer-to-peer networks in all aspects. Advantages and applications of cyber-twin in real-world environment are discussed along with the challenges the cyber-twin will face if adopted for 6G Technology. Cyber-twins are a novel and revolutionary approach in the field of Industry 4.0 which enables the escalation in the reliability or sustainability of any component.

Some topics that still need to be explored to achieve more secure, innovative, and expandable architecture are as follows:

- (1) How the cyber-twin networks can be intergraded into new business models fruitfully must be explored. The aim of usage must be defined.
- (2) Well-planned authentication framework for the cyber-twin must be developed to trace the behavior of the network.
- (3) For obtaining more efficiency and quality of services, artificial intelligence can be used for implementation and association among various cyber-twins.
- (4) Resource management is a big issue while using the clouds in the cyber-twin network which can be managed by Blockchain technology for efficient allocation of resources among various devices.

Data Availability

The data used to support the findings of this study are available upon request via e-mail: gdhiman0001@gmail.com.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] A. M. Baharudin and W. Yan, "Long-range wireless sensor networks for geo-location tracking: design and evaluation," in *Proceedings of the 2016 International Electronics Symposium, IES 2016*, pp. 76–80, Denpasar, Indonesia, September 2017.

- [2] S. Li, Y. Zhang, D. Raychaudhuri, and R. Ravindran, "A comparative study of MobilityFirst and NDN based ICN-IoT architectures," in *Proceedings of the 2014 10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, QSHINE 2014*, vol. 1, pp. 158–163, Rhodes, Greece, August 2014.
- [3] Z. Meng, Z. Chen, and Z. Guan, "Peer-to-peer file sharing in next generation eXpressive Internet Architecture," *CCF Transactions on Networking*, vol. 1, no. 1–4, pp. 28–36, 2019.
- [4] B. Gaujal, S. Haar, and J. Mairesse, "Blocking a transition in a free choice net and what it tells about its throughput," *Journal of Computer and System Sciences*, vol. 66, no. 3, pp. 515–548, 2003.
- [5] A. Albugmi, M. O. Alassafi, R. Walters, and G. Wills, "Data security in cloud computing," in *Proceedings of the 5th International Conference on Future Generation Communication Technologies, FGCT 2016*, pp. 55–59, London, United Kingdom, August 2016.
- [6] E. Pignaton, D. F. Joanna, I. Olszewska et al., "Ontological concepts for information sharing in cloud robotics," *Journal of Ambient Intelligence and Humanized Computing*, Article ID 0123456789, 2020.
- [7] Q. Yu, J. Ren, H. Zhou, and W. Zhang, "A cybertwin based network architecture for 6G," in *Proceedings of the 2nd 6G Wireless Summit 2020: Gain Edge for the 6G Era, 6G SUMMIT 2020*, Levi, Finland, March 2020.
- [8] K. Y. H. Lim, P. Zheng, and C.-H. Chen, "A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives," *Journal of Intelligent Manufacturing*, vol. 31, no. 6, pp. 1313–1337, 2020.
- [9] J. Lee, B. Bagheri, and H.-A. Kao, "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18–23, 2015.
- [10] C. Czwick and R. Anderl, "Cyber-physical twins - definition, conception and benefit," *Procedia CIRP*, vol. 90, pp. 584–588, 2020.
- [11] M. Bevilacqua, E. Bottani, F. E. Ciarapica et al., "Digital twin reference model development to prevent operators' risk in process plants," *Sustainability*, vol. 12, no. 3, pp. 1–17, 2020.
- [12] K. Agalinos, S. T. Ponis, E. Aretoulaki, G. Plakas, and O. Efthymiou, "Discrete event simulation and digital twins: review and challenges for logistics," *Procedia Manufacturing*, vol. 51, pp. 1636–1641, 2020.
- [13] A. Saad, S. Faddel, and O. Mohammed, "IoT-based digital twin for energy cyber-physical systems: design and implementation," *Energies*, vol. 13, no. 18, 2020.
- [14] F. Tao, F. Sui, A. Liu et al., "Digital twin-driven product design framework," *International Journal of Production Research*, vol. 57, no. 12, pp. 3935–3953, 2019.
- [15] Q. Yu, J. Ren, Y. Fu, Y. Li, and W. Zhang, "Cybertwin: an origin of next generation network architecture," *IEEE Wireless Communications*, vol. 26, no. 6, pp. 111–117, 2019.
- [16] S. Juneja, A. Juneja, and R. Anand, "Reliability modeling for embedded system environment compared to available software reliability growth models," in *Proceedings of the 2019 International Conference On Automation, Computational And Technology Management*, London, UK, April 2019.
- [17] A. Ahmadi, C. Cherifi, V. Cheutet, Y. Ouzrout, and I. Lyon, "Recent advancements in smart manufacturing technology for modern industrial revolution," *A Survey. Journal Of Engineering And Information Science Studies*, 2020.
- [18] W. Ge and R. Y. Zhong, "Internet of things enabled manufacturing: a review," *International Journal of Agile Systems and Management*, vol. 11, no. 2, p. 126, 2018.
- [19] K. Senthilnathan and I. Annapoorani, "Multi-port current source inverter for smart microgrid applications: a cyber physical paradigm," *Electronics*, vol. 8, no. 1, pp. 1–20, 2019.
- [20] P. K. Padhi and F. Charrua-santos, "6G enabled industrial internet of everything: towards a theoretical framework," *Applied System Innovation*, vol. 4, no. 1, pp. 1–30, 2021.
- [21] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020.
- [22] H. Park, A. Easwaran, and S. Andalarn, "Challenges in digital twin development for cyber-physical production systems," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*LNCSSpringer International Publishing, New York, NY, USA, 2019.
- [23] S. Haag and R. Anderl, "Digital twin-p," *Manufacturing Letters*, vol. 15, pp. 64–66, 2018.
- [24] J. Galvao, J. Machado, G. Prisacaru, D. Oлару, and C. Bujoreanu, "Modelling cyber-physical systems: some issues and directions," *IOP Conference Series: Materials Science and Engineering*, vol. 444, no. 4, 2018.
- [25] I. O. Olalere and O. A. Olanrewaju, "Optimising production through intelligent manufacturing," *E3S Web of Conferences*, vol. 152, pp. 3–6, 2020.
- [26] Y. Natarajan, K. Srihari, G. Dhiman et al., "An IoT and machine learning-based routing protocol for reconfigurable engineering application," *IET Communications*, pp. 1–12, 2021.
- [27] H. K. Upadhyay, S. Juneja, S. Maggu, G. Dhingra, and A. Juneja, "Multi-criteria analysis of social isolation barriers amid COVID-19 using fuzzy AHP," *World Journal of Engineering*, 2021.
- [28] A. Juneja, S. Juneja, A. Soneja, and S. Jain, "Real time object detection using CNN based single shot detector model," *Journal of Information Technology Management*, vol. 13, no. 1, pp. 62–80, 2021.

Research Article

Forecasting Variation Trends of Stocks via Multiscale Feature Fusion and Long Short-Term Memory Learning

Yezhen Liu ¹, Xilong Yu ², Yanhua Wu,² and Shuhong Song ¹

¹College of Economics and Management, Nanjing Forestry University, Nanjing 210037, China

²China Ship Scientific Research Center, China Shipbuilding Industry Corporation, Wuxi 214000, China

Correspondence should be addressed to Shuhong Song; songshuhong@njfu.edu.cn

Received 20 August 2021; Revised 8 September 2021; Accepted 11 September 2021; Published 22 September 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Yezhen Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Forecasting stock price trends accurately appears a huge challenge because the environment of stock markets is extremely stochastic and complicated. This challenge persistently motivates us to seek reliable pathways to guide stock trading. While the Long Short-Term Memory (LSTM) network has the dedicated gate structure quite suitable for the prediction based on contextual features, we propose a novel LSTM-based model. Also, we devise a multiscale convolutional feature fusion mechanism for the model to extensively exploit the contextual relationships hidden in consecutive time steps. The significance of our designed scheme is twofold. (1) Benefiting from the gate structure designed for both long- and short-term memories, our model can use the given stock history data more adaptively than traditional models, which greatly guarantees the prediction performance in financial time series (FTS) scenarios and thus profits the prediction of stock trends. (2) The multiscale convolutional feature fusion mechanism can diversify the feature representation and more extensively capture the FTS feature essence than traditional models, which fairly facilitates the generalizability. Empirical studies conducted on three classic stock history data sets, i.e., S&P 500, DJIA, and VIX, demonstrated the effectiveness and stability superiority of the suggested method against a few state-of-the-art models using multiple validity indices. For example, our method achieved the highest average directional accuracy (around 0.71) on the three employed stock data sets.

1. Introduction

Forecasting the variation trend of stocks is always one of the hot topics in the academic and practical studies of stock markets. The innately dynamic, chaotic, and nonstationary properties of stock markets make it extremely challenging to predict the tendency of financial time series (FTS) precisely. Given that the fluctuation of the stock price is affected by multiple aspects of social economic life, it has great economic and social values to forecast the developing trend of the stock price effectively. Both investors and for-profit institutions require scientific and intelligent methods to analyze and evaluate the price history so as to facilitate establishing the appropriate trade strategies.

The ultimate goal, for a specific stock, is to sell out shares at the highest price and purchase shares during the lowest period, which means minimizing risks as well as maximizing

profits. Admittedly, financial data often exhibit hybrid, nonlinear, and seemingly unrelated characteristics, which makes market hypothesis difficult to apply to predict the potential [1]. Also, some unexpected factors commonly make the stock market change dramatically, e.g., worldwide economic condition, national policies, public voices, investors' expectations, and the like.

The prevailing theory is that the stock market is largely random, especially in the case of the Iranian stock market, which is determined by certain criteria of closing price. In the past, the most traditional methods associated with time series were based on stationary trends, leading to an inherent difficulty of anticipation [2]. Therefore, plenty of researchers were devoted to conducting abundant experiments and tried to establish reliable stock price models. However, there are still massive difficulties and unsolved problems, such as those countless variables and factors. In addition, due to the

difference between short and long terms, one fine-tuning model probably works well in the short-term prediction, whereas could be poor in a longer time series.

Numerous studies have been persistently seeking suitable pathways to address such challenges, and modern artificial intelligence technologies, e.g., machine learning (ML) algorithms, have particularly facilitated this category of studies. It is a consensus that ML is qualified to extract potential characteristics and discover relative patterns from price history data. With the high-speed development of machine learning, many approaches have obtained convincing and outstanding performance on some price history data, for instance, the S&P 500 Index (S&P 500), Hang Seng Index (HSI), Jones Industrial Average (DJIA), and Nikkei 225 (N 255). Even if none of them were invariably successful in practice, their working mechanisms are worth learning and modifying, such as logistic regression, Support Vector Machine (SVM), decision tree, Recurrent Neural Network (RNN), LSTM, and Temporal Convolutional Attention-Based Network (TCAN).

In the beginning, LSTM was designed to resolve the issue of error backflow, namely, error signals explode or vanish as they flow backwards on a certain time scale. Facing noise and incompressible input sequences, LSTM can learn intervals spanning more than 1,000 iterations. A gradient-based approach is used to ensure continuous error flows in special units, which warrants that the gradient computation would be truncated at certain architecture-specific points without affecting the error flow computing on long-term data [3]. LSTM was further improved by adding the gradient propagation path and forgetting gate structure. Compared with the traditional RNN, this improvement can solve the problem of gradient vanishment. It is not the total gradient vanishment, but the one dominated by the short term that makes it difficult for models to capture long-term features. In time series prediction, such deficiency of gradient vanishment will lead to continuous loss of remote information during consecutive learning and to the dependence decrease of model parameters to remote features. Whereas stock data are exactly long-term and noncyclical and long-term dependence is critical to forecast results, the gate structure of LSTM is very suitable to complete the task of stock forecasting.

In this study, we aim to propose a novel, multiscale, convolutional feature, fusion-based LSTM model for the FTS forecasting issue. Our efforts lie in the following two aspects:

- (1) Owing to the dedicated gate structure designed for short-term and long-term memory, our model can make full use of the given stock history data to adaptively train the forecasting model. Also, the LSTM network structure effectively avoids the gradient explosion and gradient vanishment problems of deep learning. Therefore, our scheme greatly warrants the prediction performance in time series scenarios and thereby benefits the prediction of stock trends.
- (2) The multiscale, convolutional feature, fusion mechanism embedded in the proposed LSTM-based

model can diversify the feature representation and thereby can more extensively capture the feature essence of time series than traditional models. These two improvements facilitate the generalizability of our novel LSTM model to a certain extent.

The rest of the article is organized as follows. Section 2 briefly introduces the works related to stock price prediction. Section 3 illustrates our suggested methodology in details. In Section 4, experimental results and discussions are presented. Finally, we conclude the article in Section 5.

2. Related Works

Numerous studies anticipating stock price variation trends have been performed based on the FTS analysis. The widely used techniques can be roughly divided into three categories: statistical econometric models or tools, regression algorithms, and deep learning methods. This study focuses on machine learning-based techniques, so we primarily review the latter two in the following.

Timbó et al. proposed a multiple linear regression algorithm with a data processing methodology, named Knowledge Discovery in Databases (KDD) [4]. KDD is a multistep process to capture useful, vital information within massive price data sets, including selection, preprocessing, transformation, data mining, and interpretation, and in which linear regression can benefit from the precise and low-noise data. Lin et al. proposed a stock forecasting method using SVM, including two functions: feature selection and trend forecasting [5]. In terms of the technique of support vector regression (SVR), this method can forecast the tendency of stock prices well. The authors also proved the superior generalizability of this method versus others.

Compared with other machine learning methods, artificial neural networks (ANNs), particularly deep neural networks, have showed their validity in practice, such as in stock pricing prediction [6]. In the study by Wanjawa and Muchemi [7], ANN was utilized to forecast stock pricing by a feed-forward multilayer perceptron with inverse loss propagation and thus obtained good performance. Nonetheless, researchers noticed that ANN scarcely establishes a correlative connection between current and previous data, leading to poor robustness and low universality. It is a consensus that the predicted results are correlative not only to current data but also to previous data. To tackle this problem, RNN was devised. In contrast to traditional ANN, RNN proved more convincing performance in the financial field. During its iterations, earlier time series data are beneficial to the model's precise learning via the feed-forward and back-forward looping.

The earlier stock prediction used traditional RNNs. These methods were frequently combined with other technologies for denoising data, such as Discrete Wavelet Transform (DWT). Contrary to the limitations of Fourier Transform, DWT originally uses wavelet basis to describe the signal [8]. The wavelet basis is a very small scale of a signal, so the wavelet transform has the ability to describe time series. One of the keynotes of wavelet transform is to use different resources to

describe different frequency ranges. For different frequency scales, like trees, the richness of the description is very different. The higher the frequency of sampling, the finer the description. Compared with DWT, b-Spline Wavelets of High-Order-d (BSd) can achieve better results on certain data sets. BSd-RNN was proposed to forecast the high-frequency time series in the study by Hajiabotorabi et al. [9]. With the combination of BSd and RNN, the time series was decomposed into numerous smooth data sets using a multiresolution technique, which made it possible to generate distinctly detailed data sets with modest wave amplitudes. Due to the local properties, the suggested BSd-RNN model was capable of accurately approximating more smooth patterns than other common models.

Among all RNN-based models, LSTM could be the most effective model for time series prediction. LSTM used a set of memory cells with the gate structure to replace hidden neurons of RNN. As such, through the gate structure feature, the information was retained and persistently updated in the following training iterations. For instance, Zhao et al. used LSTM to achieve extraordinary performance in stock trend prediction [10]. Seng et al. used the ordinary three-layer LSTM structure, instead of utilizing too complex network structures, to forecast LQ45 financial sectors indices and obtain nice results [11].

Also, many researchers sought the manners to further improve LSTM's performance by assembling other learning models. Autoregressive Fractional Integrated Moving Average (ARFIMA) was first used to predict the weather's seasonal change. Afterwards, some researchers attempted to forecast stock prices using ARFIMA because it is fairly suitable for predicting the results of time series data. ARIMA-LSTM model was proposed by Bukhari et al. [12]. In the field of deep learning, Convolution Neural Network (CNN) is another outstanding network qualified for forecasting tasks using varied convolution blocks. Nonetheless, CNN can hardly tackle time series data separately. Thus, Qiu et al. combined RNN with CNN to put forward a novel network called Deep Wide Area Neutral Network (DWNN) [13]. Experiments showed that this model can reduce the mean square error of prediction by 30% compared with the conventional RNN structures. In order to capture the time-dependent characteristics, Zhang et al. proposed another alternative fusion of CNN and RNN strategy in [14]. During the hidden state transfer, CNN's convolution layers were introduced to extract the correlation features and RNN, meanwhile, proceeded in time steps. As such, this design had not only the depth of RNN in the temporal dimension but also the width of temporal data.

One of the major drawbacks of aforementioned methods is their inability to forecast highly dynamic and transforming patterns of stock price variation, whereas TCAN solved this issue to a certain extent. Hao et al. proposed the TCAN algorithm that integrated an attention mechanism into the time series neural network [15]. By jointly introducing Temporal Attention (TA) and Enhanced Residual (ER), TCAN was enabled to extract both the shallow layers' pivotal features and correlative characteristics hidden in the time series.

3. Methodology

In this section, we detail our novel LSTM-based structure and scheme for stock price prediction as follows.

As shown in Figure 1, our method is mainly composed of two parts: data preprocessing and model construction. First, the raw data set is preprocessed with wavelet denoising, normalized time step data, and data set division. For the latter, the prediction model employs the three-branch structure to constitute the multiscale feature fusion-based convolutional LSTMs, followed by a dense layer for eventually denormalizing the output.

3.1. Preprocessing. The Yahoo Finance data (<https://finance.yahoo.com>) was downloaded to act as our experimental data sets, which includes the data of open, high, low, and closing prices of stocks, trading volume, and adjusted prices.

To capture the essential characteristics of stock price time series data, besides, the six originally contained variables (i.e., *open*, *closing*, *high*, *low*, *trading volume*, and *adjusted price*), the *moving average* (MA), and *exponent moving average* (EMA) are also calculated in our study. These two can reflect the trend of stock price variation exponentially or at a constant level and are proved effective to guide stock investment. In this way, the stock price history data are represented as the form of eight-dimensional time series and further used as the input for our LSTM-based model.

The complexity and volatility of the stock market and the dynamic trading criterion usually cause the stock price data obtained to be noisy [16] and nonstationary because enormous factors, either explicitly or implicitly, influence the variation tendency of stock prices. Classic denoising algorithms, such as Fourier analysis, are prone to being ineffective in the case of massive information fusion. Therefore, we employ the wavelet transform to denoise the financial time series data. The wavelet transform has the capacity to conduct time series analysis in both the time and frequency domains. Specifically, the *db4* wavelet function, having four decomposition layers, is used to remove the noise hidden in high frequencies.

It is explicit that single stock price data cannot reflect the tendency of stock price variation and are not qualified to forecast the future pricing. Therefore, the data utilized in our model are extracted from the whole data set at all time steps. The time interval is set to 20 days in our current study. After the whole preprocessing, the sequential data would be represented as a $b \times t \times d$ matrix, in which b , t , and d represent the batch size, time step number, and feature dimensionality of the input data, respectively.

Among all the eight adopted FTS variables, the closing price could be the most straightforward one because it impacts investment strategies to a great extent, and thus, the closing price is regarded as one primary prediction target in our study.

3.2. The Proposed LSTM-Based Model Structure. The chaotic, nonstationary, and nonlinear characteristics of stock pricing limit the feasibility of conventional neural networks. RNN

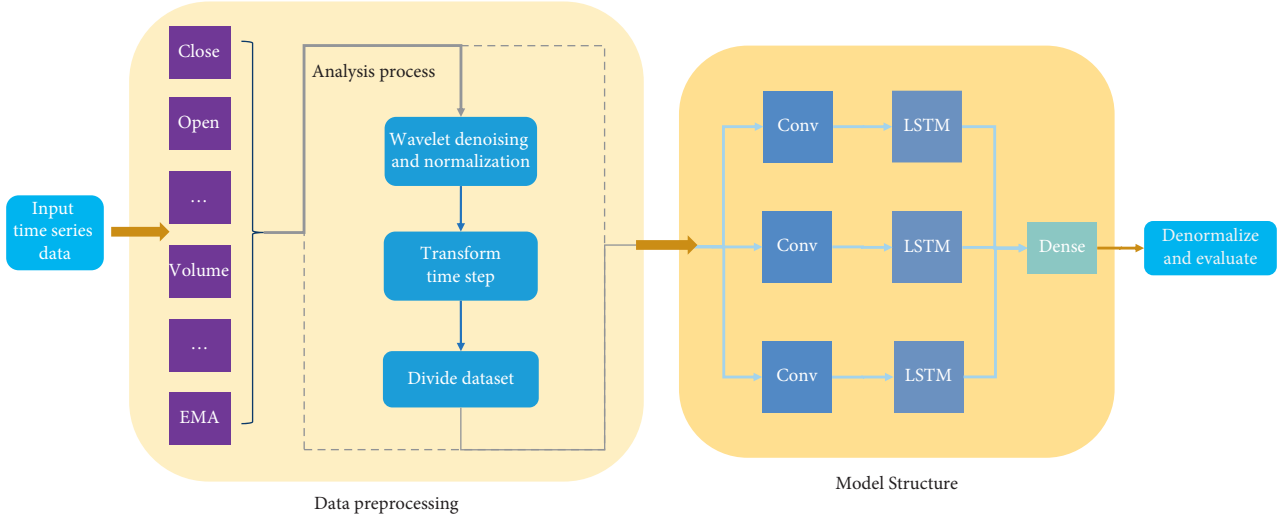


FIGURE 1: Scheme of our FTS prediction model.

was designed for avoiding long-term dependence problems and using the correlated information between various time intervals when tackling the time series prediction problem. However, original RNN can hardly conquer the challenge of gradient vanishment. Different from classic RNN, the LSTM neural network consisting of special memory cells was proposed by Abedinia et al. [17]. In LSTM, the memory cells are used to replace the hidden layer neurons in RNN, and the states of memory cells play the pivotal roles in the forecasting task. Moreover, the gate structures are utilized to convey the feature information and to update the state of memory cells. As shown in Figure 2, each LSTM cell is composed of three dynamic gates: the input, forget, and output gates. The basic structure of any memory cell includes one add layer, two tanh layers, three sigmoid layers, and three concatenation layers.

The forget gate determines how much cell state information would be discarded from the previous cell. As shown in Figure 2, one memory cell accepts the output of previous information h_{t-1} and external input of x_t in a concatenated vector $[h_{t-1}, x_t]$ via σ transformation, as listed in Equation (1) in which W_f and b_f separately represent the weight matrix and bias of the forget gate. After Equation (1) calculation, f_t ranging from 0 to 1 determines the reserved percentage of the previous cell state C_{t-1} , where 0 indicates the entire abandonment and 1 indicates the entire acceptance.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (1)$$

In LSTM's cell model, the input gate can determine the proportion of the new input x_t reserved to generate the eventual cell state C_t . This gate extracts pivotal information from current input as well as prevents unconsidered content from entering current cell. The calculation of the input gate is detailed in Equation (2), where W_i and b_i are separately the weight matrix and bias.

The updated information of cell state \tilde{C}_t is generated through the tanh layer and using Equation (3) in which

W_c and b_c denote the weight matrix and bias, respectively. The current cell state C_t can be obtained using Equation (4).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2)$$

$$\tilde{C}_t = \tan h(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t. \quad (4)$$

The output gate determines how much cell state C_t can be transformed into the output h_t using Equation (6).

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (5)$$

$$h_t = O_t \times \tan h(C_t). \quad (6)$$

In our empirical studies, the final output of our designed model is the predicted value of the $(t+1)$ -th day's closing stock price using the previous t days' feature data. Mean square error (MSE) measurement is used to constitute the ultimate loss function of our LSTM-based model.

Convolutional LSTM (ConvLSTM) [18] has proven the excellent performance while participating in time series prediction problems because it is good at capturing the spatio-temporal relations well. Moreover, the multiscale feature fusion strategy overall outperforms other basic structures on deep feature extraction. Therefore, to further improve the time series feature extraction effectiveness, in our LSTM-based model (see Figure 1), we devise a multiscale convolutional feature fusion mechanism to extensively extract the features of stock pricing history, i.e., the three-branch structure on the right in Figure 1. However, due to the characteristics of stock pricing data, one-dimensional convolutions are used in our model. It is worth mentioning that the three convolutional layers use the uniform number of filters (e.g., 100), whereas the kernel sizes and strides are set differently, e.g., 6 (kernel size) and 3 (stride) for the first convolutional layer, 12 (kernel size) and 3 (stride) for the second, and 6 (kernel size) and 2 (stride) for the third. As such,

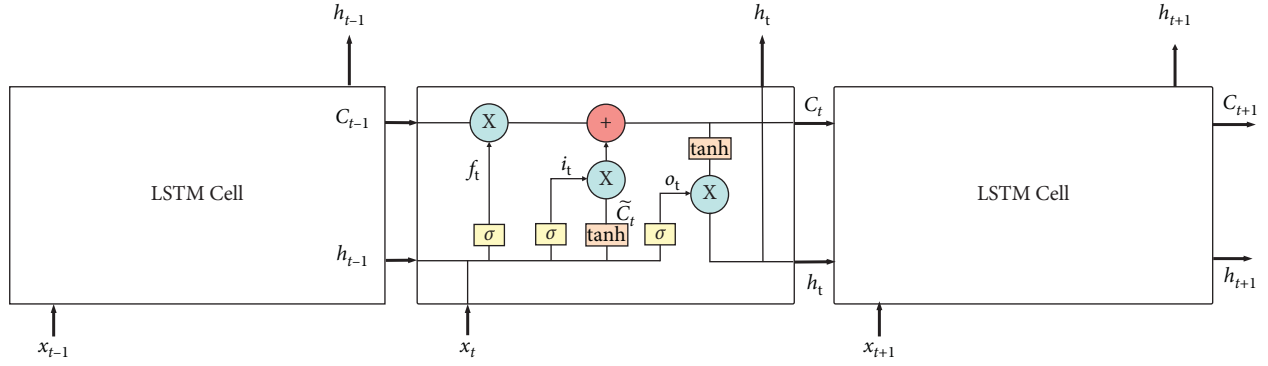


FIGURE 2: Gate structure of LSTM.

the three convolutional layers input different scales of features into the following LSTMs, and we achieve the goal of FTS feature presentation diversity. Finally, the outputs of the three branches are concatenated as the input of the Dense layer.

3.3. Data Set. As mentioned in Section 3.1, we conducted experiments on three influential stock index data sets, i.e., S&P 500, DJIA and CBOE Volatility Index (VIX). To obtain relatively low-noisy data, the experimental data sets were constituted by discrete time series data coming from the three stock history data sets. As previously introduced, eight FTS variables are used for model training and forecasting, i.e., open, closing, high, low, trading volume, adjusted price, MA, and EMA. Among them, the former six are from the original data sets, and the last two are regenerated using the following equations:

Let

$$\text{Avg } M_i = \frac{1}{N} \sum_{i=N}^i (\text{close}_i), \quad (7)$$

where $\text{Avg } M_i$ and close_i represent i -th day's moving average and closing price, respectively, and N denote the time step length. Then,

$$MA = \ln \frac{\text{close}}{\text{Avg } M}. \quad (8)$$

Let

$$E_i = \frac{2(\text{close}_i - E_{i-1})}{N + 1} + E_{i-1}, \quad (9)$$

where E_i represents i -th day's exponent moving average; then,

$$EMA = \ln \frac{\text{close}}{E}. \quad (10)$$

4. Experimental Studies

4.1. Setup. To evaluate the realistic performance of our devisal, three well-established machine learning algorithms were adopted to make comparisons with our proposed LSTM-based model, including ANN, SVR, and linear

regression. Besides, five validity indices were used for performance measurement: mean square error (MSE), mean absolute percentage error (MAPE), mean absolute error (MAE), coefficient of determination (R^2), and directional accuracy (DA). Their detailed definitions are as follows:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2, \quad (11)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|, \quad (12)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{|\hat{y}_i - y_i|}{y_i} \right|, \quad (13)$$

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^N (\hat{y}_i - \bar{y}_i)^2}, \quad (14)$$

$$DA = \frac{1}{n} \sum_{i=1}^n a_i, \quad (15)$$

in which N represents the sample size, y_i represents the truth value, \bar{y}_i represents the mean value of the truth value, and \hat{y}_i represents the predicted value. a_i in Eq. (15) signifies whether the rising and falling forecasts are correct, and 0 for false and 1 for true.

As mentioned in Section 3, for the three input convolutional layers in Figure 1, we uniformly set the filter number to 100, employing different kernel sizes and strides. Specifically, the upper convolutional layer has the kernel size 6 and stride 3, the medium layer has the kernel size 12 and stride 3, and the bottom layer has the kernel size 6 and stride 2. The Leaky ReLU was used as the activation function in these convolutional layers.

Our model was trained by the Adam optimizer with an initial learning rate of $2e-6$, which drops every five iterations at 0.95. Meanwhile, three adjacent LSTM layers were equipped with 128 units and 0.2 dropout rate.

In addition, ANN was also trained using the Adam optimizer, consisting of five dense layers with units 500, 500, 250, 250, and 1, respectively. SVR employed the radial basis function (RBF) as the kernel function.

TABLE 1: Experiment of different time step lengths of LSTM on S&P 500.

Time step	MSE	MAE	MAPE	R^2
5	$4.9663e-3$	$4.6879e-3$	$3.7178e-3$	0.9747
10	$6.3305e-5$	$5.7630e-3$	$4.4984e-3$	0.9804
20	$3.7410e-5$	$4.1263e-3$	$3.2351e-3$	0.9645
30	$2.2613e-5$	$3.6213e-3$	$2.6458e-3$	1.0343
40	$1.4465e-4$	$9.2449e-3$	$7.2079e-3$	0.9133

TABLE 2: Performance comparisons among four methods on S&P 500.

Method	MSE	MAE	MAPE	R^2
Proposed LSTM-based model	$2.2613e-5$	$3.6213e-3$	$2.6458e-3$	1.0343
ANN	$2.9482e-5$	$4.7260e-3$	$1.6978e-3$	1.0299
SVR	$4.4264e-3$	$6.5222e-2$	$4.7541e-3$	0.7755
Linear regression	$1.0189e-4$	$9.9822e-3$	$7.2890e-3$	1.0087

TABLE 3: Performance comparisons among four methods on DJIA.

Method	MSE	MAE	MAPE	R^2
Proposed LSTM-based model	$1.2958e-3$	$8.8381e-3$	$5.9901e-3$	0.9721
ANN	$5.1614e-4$	$9.9758e-3$	$6.7933e-3$	0.9605
SVR	$3.7029e-3$	$5.6042e-2$	$3.7629e-2$	0.7696
Linear regression	$1.0295e-4$	$9.3711e-3$	$6.3274e-3$	0.9899

TABLE 4: Performance comparisons among four methods on VIX.

Method	MSE	MAE	MAPE	R^2
Proposed LSTM-based model	$3.0961e-5$	$3.8656e-3$	$9.9029e-3$	0.9958
ANN	$1.4427e-4$	$2.0201e-2$	$6.4719e-2$	1.0222
SVR	$2.3282e-3$	$4.3949e-2$	$1.2084e-1$	0.9218
Linear regression	$8.7053e-5$	$6.0373e-3$	$1.5141e-2$	0.9736

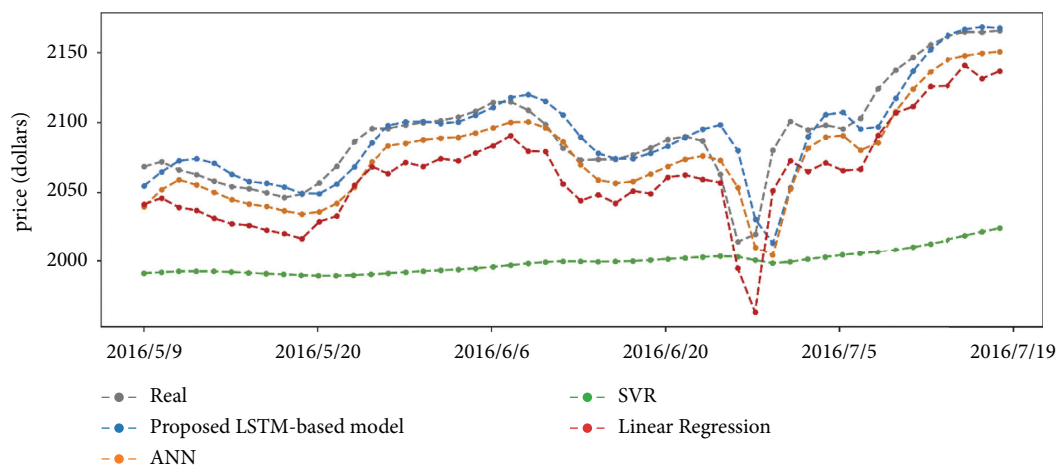


FIGURE 3: Comparison of predicted closing prices of four methods on the S&P 500 data set.

4.2. *Experimental Results and Analyses.* The time step length of time series is usually a core parameter determining FTS forecasting efficiency. Table 1 shows the relationships of various time step numbers with LSTM's realistic performance. Usually, overlong time step lengths are prone to the

gradient vanishment, whereas too short ones easily lose the vital information embedded in time series. As revealed in Table 1, the step-length 5 obtained the worst score, which implies that it is too short for FTS forecasting tasks, whereas the step-length 40 seemed overfitting a bit. Generally, he

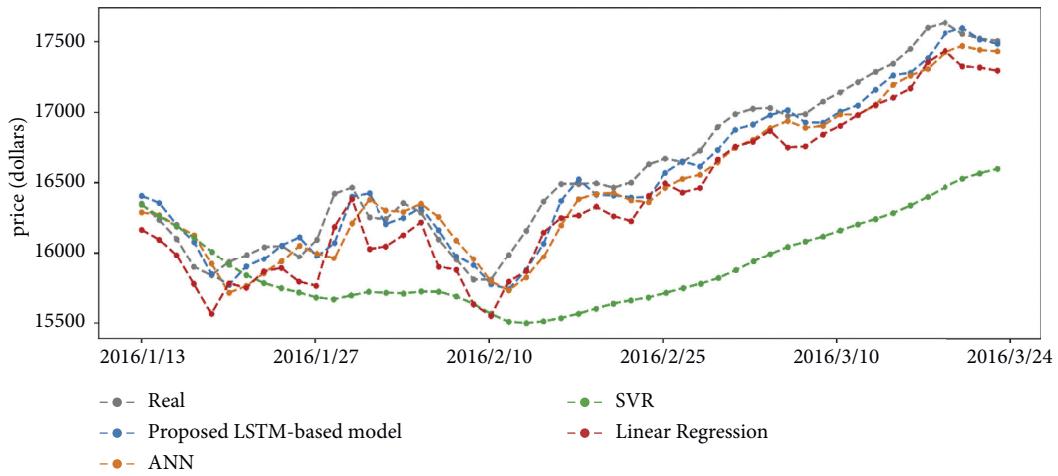


FIGURE 4: Comparison of predicted closing prices of four methods on the DJIA data set.

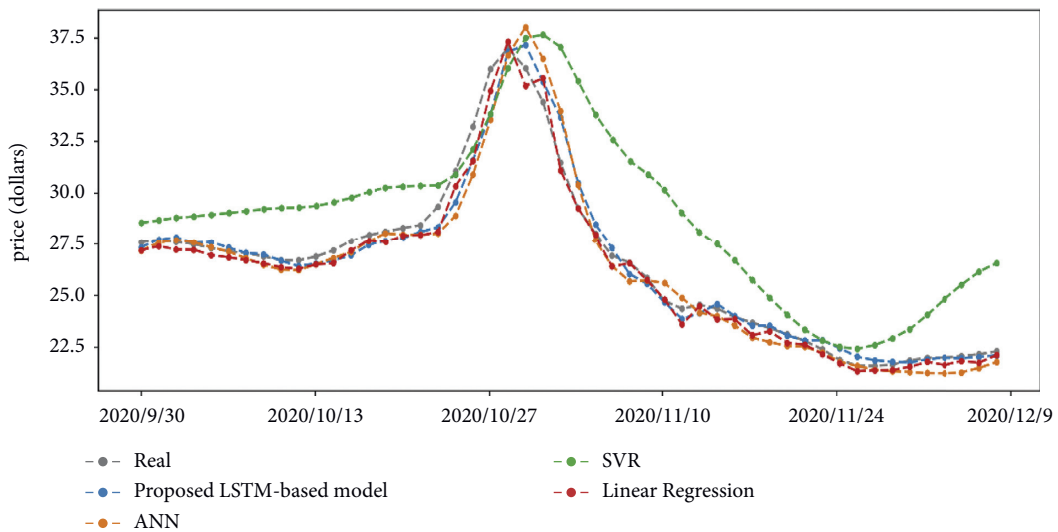


FIGURE 5: Comparison of predicted closing prices of four methods on the VIX data set.

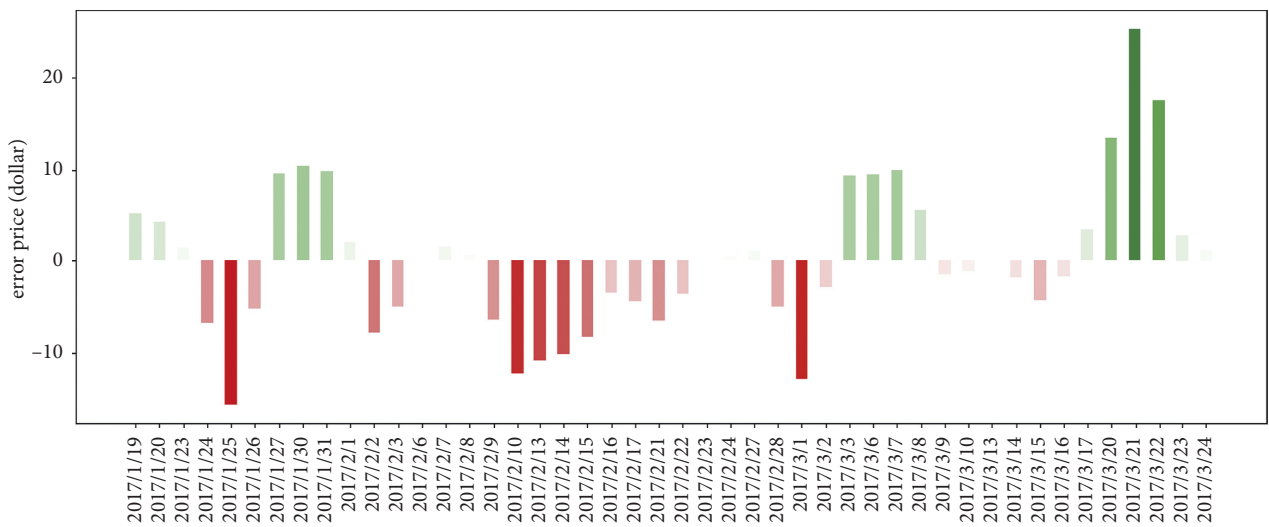


FIGURE 6: Difference between prediction and ground truth of our method on the S&P 500 data set.

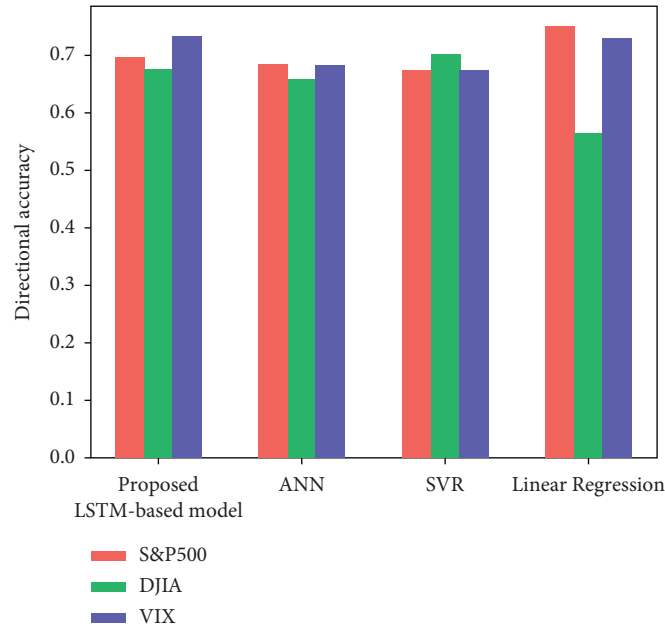


FIGURE 7: Comparisons of DA index of different methods.

step-length 30 obtained the best performance. Thus, we set the time step length to 30 in our empirical studies.

Tables 2–4 specifically display the obtained scores of the four methods on three stock data sets by means of used validity indices. Generally, our LSTM-based model scores best on all indices, particularly on S&P 500 and VIX data set. Despite that the results of our method on DJIA are not overwhelming, they still rank at top 2.

Figures 3–5 further intuitively show the predicted trends of closing prices on three stock data sets. As is revealed, the results of the proposed LSTM-based method are closer to the real market trends than those of the others, as overall the blue lines have the universality of the smallest offset/deviation from the ground truth. By the way, some blue lines even overlap with the grey lines at some points, which implies the preferable sensitivity to small variations of our designed, multiscale convolutional feature fusion-based LSTM model.

Because the stock price is almost noncyclical, it is reasonable that the predicted results of all employed methods have the characteristic of hysteresis. The adoption of wavelet transform can lighten such influence to a certain extent by removing noise as well as retaining inherent features; however, it is not enough. In our proposed method, with embedding the multiscale feature fusion mechanism, it is distinct that the hysteresis of our method is less than that of ANN. As the evidence, in Figure 3, the prediction curve of the ANN is approximately the back translation of our LSTM-based method. Figures 3–5 also illustrate that the overall accuracy and stability of our method are better than those of the other methods.

Figure 6 displays the bias between our LSTM-based model's predicted closing prices and the given closing prices on S&P 500 data set. In this figure, the green bar represents the case where the prediction is higher than the truth,

whereas the red represents the inverse case. As is shown, the highest bias is less than 30 and generally varies between -10 and 10. Compared with the stock price high up to 2,500, these deviations actually reflect the forecasting stability and effectiveness of our method.

The DA index can reflect whether the forecasted trend conforms with the real movement tendency of closing prices of stocks. The rising or falling tendency is another remarkable indicator in stock trading. Hence, we utilized DA to prove the superiority of our efforts, as shown in Figure 7. We also achieve the similar conclusion that overall, our LSTM-based model has the higher prediction accuracy on stock tendency changes.

All the above results and analyses indicate that the dedicated gate structure of LSTM as well as the proposed multiscale feature fusion strategy greatly warrant the desirable preferable performance of our method in forecasting stock trends.

5. Conclusion

In this paper, we propose a multiscale convolutional feature fusion-based LSTM model to address the challenge of forecasting stock trends. With experiments on three classic stock data sets, it has been proved that the proposed method has superior effectiveness and stability than a few other state-of-the-art methods. For future study, we will contribute to further improving the prediction accuracy based on other deep learning techniques.

Data Availability

The labeled data sets 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



This work was supported in part by Nanjing Forestry University.

References

- [1] S. Akhter and M. A. Misir, "Capital markets efficiency: evidence from the emerging capital market with particular reference to Dhaka stock exchange," *South Asian Journal of Management*, vol. 12, pp. 35–51, 2005.
- [2] M. Nabipour, P. Nayyeri, H. Jabani, and A. Mosavi, "Deep learning for stock market prediction," *Entropy*, vol. 22, 2020.
- [3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, 1997.
- [4] N. S. Timbo, S. Labidi, T. P. d. Nascimento, M. L. Lima, G. Nunes Neto, and R. C. Matos, "Approach based on linear regression for stock exchange prediction—case study of Petr4 petrobrás, Brazil," *International Journal of Artificial Intelligence & Applications*, vol. 7, no. 1, pp. 21–31, 2016.
- [5] Y. Lin, H. Guo, and J. Hu, "An svm-based approach for stock market trend prediction," in *Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN)*, Dallas, TX, USA, August 2013.
- [6] P. Yu and X. Yan, "Stock price prediction based on deep neural networks," *Neural Computing & Applications*, vol. 32, no. 6, pp. 1609–1628, 2020.
- [7] B. W. Wanjawa and L. Muchemi, "Ann model to predict stock prices at stock exchange markets," 2014, <https://arxiv.org/abs/1502.06434>.
- [8] L. Bai, S. Yan, X. Zheng, and B. M. Chen, "Market turning points forecasting using wavelet analysis," *Physica A Statistical Mechanics & Its Applications*, vol. 437, pp. 184–197, 2015.
- [9] Z. Hajiabotorabi, A. Kazemi, F. F. Samavati, and F. M. M. Ghainia, "Improving DWT-RNN model via B-spline wavelet multiresolution to forecast a high-frequency time series," *Expert Systems with Applications*, vol. 138, Article ID 112842, 2019.
- [10] Z. Zhao, R. Rao, S. Tu, and J. Shi, "Time-weighted LSTM model with redefined labeling for stock trend prediction," in *Proceedings of the IEEE International Conference on Tools with Artificial Intelligence*, Boston, MA, USA, November 2017.
- [11] H. Seng and J. C. Young, "Predicting LQ45 financial sector indices using RNN-LSTM," *Journal of Big Data*, vol. 8, 2021.
- [12] A. H. Bukhari, M. Raja, M. Sulaiman, S. Islam, and P. Kumam, "Fractional neuro-sequential arfimalstm for financial market forecasting," *IEEE Access*, vol. 8, p. 1, 2020.
- [13] J. Qiu, B. Wang, and C. Zhou, "Forecasting stock prices with long-short term memory neural network based on attention mechanism," *PLoS One*, vol. 15, no. 1, Article ID e0227222, 2020.
- [14] R. Zhang, Z. Yuan, and X. Shao, "A new combined CNN-RNN model for sector stock price analysis," in *Proceedings of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, Tokyo, Japan, July 2018.
- [15] H. Hao, Y. Wang, Y. Xia, J. Zhao, and F. Shen, "Temporal convolutional attention-based network for sequence modeling," 2020, <https://arxiv.org/abs/2002.12530>.
- [16] O. Dessaint, T. Foucault, L. Frésard, and A. Matray, *Noisy Stock Prices and Corporate Investment*, Social Science Electronic Publishing, Rochester, NY, USA, 2019.
- [17] O. Abedinia, N. Amjadi, and H. Zareipour, "A new feature selection technique for load and price forecast of electrical power systems," *IEEE Transactions on Power Systems*, vol. 32, p. 1, 2016.
- [18] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong, and W. C. Woo, "Convolutional LSTM network: a machine learning approach for precipitation nowcasting," in *Proceedings of the 28th International Conference on Neural Information Processing Systems*, Montreal, Canada, December 2015.

Research Article

Material Discrimination Algorithm Based on Hyperspectral Image

Jian Zhou ¹, Zhuping Wang,² Yingjie Jiao,¹ and Cong Nie ¹

¹*Xi'an Modern Control Technology Research Institute, Xi'an, China*

²*Xi'an Institute of Electromechanical Information Technology, Xi'an, China*

Correspondence should be addressed to Cong Nie; abbqq_15986@163.com

Received 16 July 2021; Revised 18 August 2021; Accepted 24 August 2021; Published 13 September 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Jian Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hyperspectral information can be used to express the material properties of objects, which has a strong effect on camouflage recognition. However, it is difficult to process it directly because of the huge hyperspectral image data. Therefore, this paper proposes a new band selection algorithm to achieve band selection by simulating visual perception. The subspace clustering self-attention adversarial network is constructed to realize the initial selection of band. According to the visual chromatic aberration principle, a model is constructed to determine the band that combines the strongest response intensity of a particular material, and then this band is selected as the final band, therefore realizing the algorithm of material demarcation in this way.

1. Introduction

Visible near-infrared band images are obtained by sensors through detecting the electromagnetic radiation reflection of objects. It can precisely characterize ground objects so that each object has a spectral fingerprint which is of great significance to the identification of object materials [1, 2]. However, a hyperspectral image has high spectral dimension and spatial resolution, so it is difficult to process it directly because of a large amount of data [3, 4]. Thus, more in-depth studies have been carried out: in 2010, Yang et al. [5] used a supervised way to select band signals; Di et al. [6] applied a band selection to human face recognition and achieved good results. In 2010, Li and Qian [7] constructed a sparse matrix to analyze different bands; Samadzadegan and Mahmoudi [8] constructed the swarm intelligence to optimize band selection strategy. In 2012, Du et al. [9] established a collaborative sparse model to select hyperspectral bands; Hedjam and Cheriet [10] realized a band selection based on graph clustering. In 2013, Feng et al. [11] realized the band selection based on trivariate mutual information and clonal selection. Nakamura et al. [12] proposed a nature-inspired framework for band selection. In 2014, Su et al. [13] used the particle swarm optimization to optimize the band selection process; Xiurui Geng et al. [14] realized a band selection

through gradient analysis of different band images. In 2015, Jia et al. [15] proposed a band selection scheme based on the idea of sorting; Patra et al. [16] introduced the idea of rough set to select bands. In 2016, Feng et al. [17] utilized the multiple kernel learning based on discriminative kernel clustering for hyperspectral band selection; Liu et al. [18] proposed a band selection algorithm based on the distribution of adjacent pixels. In 2017, Cao et al. [19] improved a classification map algorithm for fast hyperspectral selection; Shah et al. [20] proposed an algorithm of the dynamic frequency domain to realize band selection. In 2018, Wang et al. [21] proposed the optimal clustering framework to achieve hyperspectral band selection; Xie et al. [22] made modeling and analysis according to the representativeness of the bands. In 2019, Sun et al. [23] used a weighted kernel regulation to realize band selection. Sun et al. [24] calculated the variance between spectral bands and built a model for band selection. In 2020, Torres et al. [25] applied a band selection into the field of signal enhancement. Sun et al. [26] used the idea of low rank to cluster hyperspectral bands. Patra and Barman [27] focused on the image boundary intensity to realize band selection based on the fuzzy set.

To sum up, main problems of hyperspectral band selection are as follows. (1) It is difficult to establish a unified band selection model due to high dimensions of

hyperspectral data. (2) The quality of band selection cannot directly show its effect. Therefore, according to the above problems, (1) a hyperspectral band selection algorithm is constructed based on vision and (2) a subspace clustering framework is proposed based on deep adversary for realizing the preliminary clustering of spectral information. A perception model based on color is proposed to visualize the difference between the target and the background to show the perception effect.

2. A Visual Perception Algorithm

More than 80% information is obtained by human vision. One object can be recognized and distinguished from the background mainly by the color. At present, the captured natural images can be regarded as the superposition of RGB three channels. On this basis, a large number of research studies on target extraction, image retrieval, and analysis have been carried out, and a series of achievements have been achieved. The natural images can be regarded as hyperspectral data with low number of channels. Therefore, we migrate the related algorithms of RGB images to the hyperspectral field, and the process is shown in Figure 1. (1) The subspace clustering network of deep countermeasure is constructed to realize the initial band selection. (2) According to the color difference of vision principle, a model is constructed to determine the band combination with the strongest response intensity of a specific material, and then the band is selected as the final band.

2.1. Subspace Clustering Based on Self-Attention Adversarial. When high-dimensional data are encoded to output low-dimensional feature representation, a large amount of information will be lost. However, the attention model, which is based on encoder-decoder framework, can lose less information.

$$\text{Attention}(Q, K, V) = s(Q, K^T)V, \quad (1)$$

where $S(\cdot)$ is the similarity function and Q is the output information. The self-attention adversarial model we built is structured as Figure 2.

For true sample acquisition, k groups of $A = \{A_1, A_2, \dots, A_k\}$ are obtained through the similarity matrix learned by word expression layer. the projection residual from A_i to the corresponding subspace S_i is calculated as follows:

$$L_r(Z_i) = \|Z_i - V_i V_i^T Z_i\|_2^2, \quad (2)$$

where Z is the characteristic matrix, V_i is the projection matrix, T represents matrix transpose, and L_R represents the projection residual. m data with small residuals are selected as positive samples. The corresponding generator resistance loss function is as follows:

$$L_g = \frac{1}{km} \sum_{i=1}^k \sum_{j=1}^m L_r(\bar{Z}_j). \quad (3)$$

For false samples, the sampling layer randomly samples from the estimated subspace S_i to generate m false samples $\bar{Z}_j = \sum \theta_j Z_j$. In order to make the generated data closer to the learning subspace of discriminator, the antiloss is introduced to revise the existing loss function:

$$L_G(C) = L_g + \frac{\lambda}{2} \|X - \hat{X}\|^2 + \frac{\alpha}{2} \|U - UC\|^2 + \frac{\beta}{2} \|C\|_p, \quad (4)$$

where λ is the balanced sparse.

A discriminator is constructed by projection residuals to distinguish true and false samples, and the probability loss function of samples belonging to subspace is established.

$$L_d = \frac{1}{m} \sum_{j=1}^m \{L_r(Z_j) + \max(0, \varepsilon - L_r(Z_j))\}, \quad (5)$$

where ε is the parameter. The loss function of the discriminator corresponding to k -rent is as follows:

$$L_d = \min_{V_1, V_2, \dots, V_k} \frac{1}{k} \sum_{i=1}^k L_d. \quad (6)$$

The second term of the following formula is introduced to increase the separation of different groups of subspaces after introducing regular terms. The third term of the following formula is to reduce V_i :

$$L_E = L_d + \mu_1 \sum_{i \neq j} \|V_i^T V_j\|_F^2 + \mu_2 \sum_{i=1}^k \|V_i^T V_i - I\|_F^2, \quad (7)$$

where μ_1 and μ_2 are two constants greater than 0.

In order to make better use of the local manifold structure information of the image, Laplacian regularization term is introduced into the loss function of the generator to construct the image connection relationship. The weight can be expressed as follows:

$$W_{ij} = \begin{cases} \exp\left(\frac{\|x_i - x_j\|}{2\sigma^2}\right), & x_i, x_j \in N_k, \\ 0, & \text{others,} \end{cases} \quad (8)$$

where N_k is the K neighborhood of n vertices. For the nonlinear manifold structure, the energy function is defined as follows:

$$Q = \min_q \sum_{i,j=1}^n \|q(x_i) - q(x_j)\|^2 W_{ij}. \quad (9)$$

According to the definition of the Laplace matrix,

$$L = D - W \left(\text{s.t. } D_{ii} = \sum_j W_{ij} \right). \quad (10)$$

Q can be rewritten as

$$Q = \min_q \text{tr}(\bar{Q} L \bar{Q}^T), \quad (11)$$

$$\bar{Q} = [q(x_1), q(x_2), \dots, q(x_n)].$$

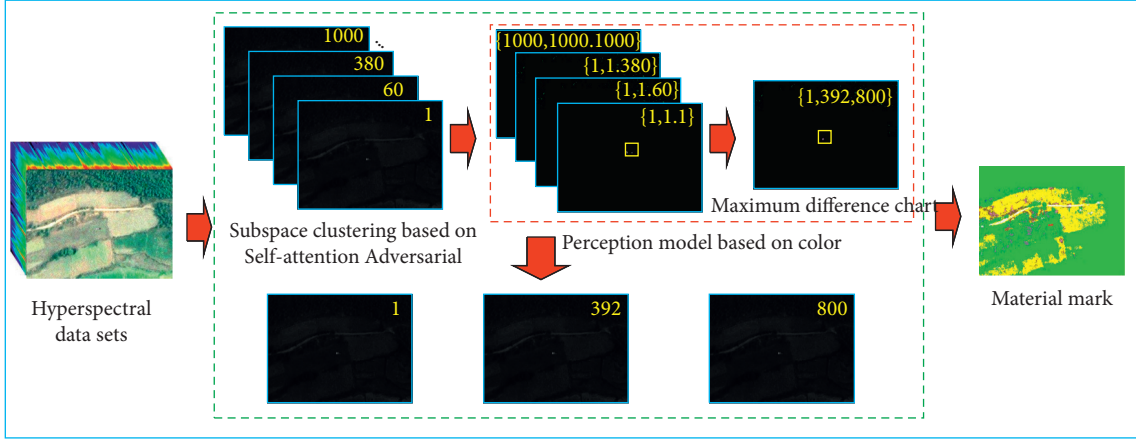


FIGURE 1: Flow chart of the band selection algorithm.

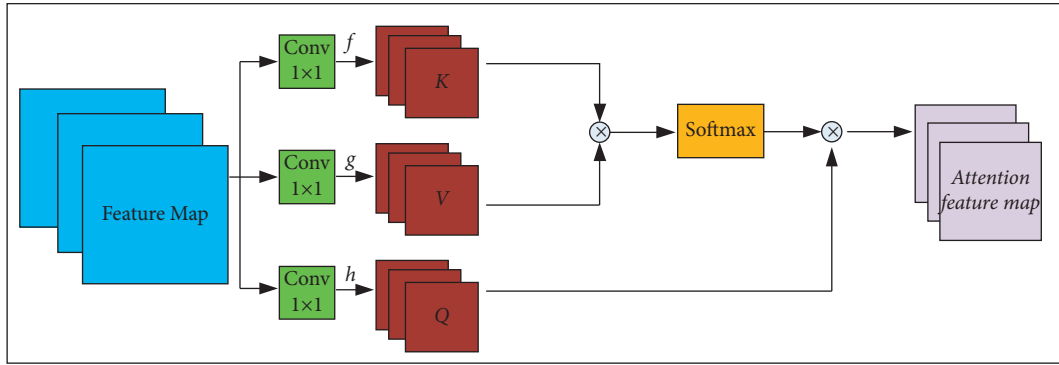


FIGURE 2: Self-attention adversarial network.

The final generator loss function is obtained as follows:

$$L_G(C) = L_g + \frac{1}{2} \{ \lambda \|X - \hat{X}\|^2 + \alpha \|U - UC\|^2 + \beta \|C\|_p + \gamma \text{tr}(CLC^T) \}. \quad (12)$$

2.2. Perception Model Based on Color. According to Gestalt psychocognitive analysis, objects can be recognized, mainly by the eyes and brain. When the eyes observe the images, they cluster themselves according to certain rules to make them become a comprehensible structural entity. Among them, color feature is an effective way.

On the basis of the previous analysis in the last section, in order to more comprehensively express the spectral information between bands, a model is established on the basis of relative entropy:

$$S(X, Y) = D_{KL}(P|Q) - \lambda M_t(X, Y),$$

$$D_{KL}(P|Q) = \sum P(x) \ln \frac{P(x)}{Q(x)}, \quad P(x), Q(x) \in [0, 1],$$

$$M_t = \begin{bmatrix} m(X_1, X_1) & \dots & m(X_1, X_M) \\ \dots & \dots & \dots \\ m(X_M, X_1) & \dots & m(X_M, X_M) \end{bmatrix}, \quad (13)$$

$$m(X, Y) = \sum_{i=1}^n \sum_{j=1}^m P(x_i y_j) \log_2 \frac{P(x_i y_j)}{P(x_i)},$$

where $S(X, Y)$ is the spectral correlation between band X and band Y ; $D_{KL}(P|Q)$ is the relative entropy of probability distribution corresponding to band X and band Y ; and $M_t(X, Y)$ is the average mutual information between band X and band Y . λ is the weight coefficient, which is determined by the relative amplitude of $D_{KL}(P|Q)$ and $M_t(X, Y)$.

In order to better integrate the entropy information into it, we modify $S(X, Y)$. Because the entropy divergence and the average mutual information are basically equally important to the band selection process, in order to make the contribution of the corresponding matrix M_{KL} and M_t consistent to S , the normalization function is constructed.

$$S(X, Y) = C \times M_{KL} - M_t,$$

$$M_{KL} = \begin{bmatrix} 0 & \dots & D_{KL}(X_1|X_M) \\ & & \dots \\ D_{KL}(X_M|X_1) & \dots & 0 \end{bmatrix}, \quad (14)$$

$$C = \frac{\sum_{i=1}^M \sum_{j=1}^M |M_t(i, j)|}{\sum_{i=1}^M \sum_{j=1}^M |M_{KL}(i, j)|}$$

Different pixel values may correspond to the same color name. For this purpose, we construct the mapping relation; given the data $D = \{d_1, \dots, d_N\}$, the corresponding word is $W = \{w_1, \dots, w_M\}$; these words are considered from the potential theme $Z = \{z_1, \dots, z_K\}$. Therefore, a probability model is constructed as follows:

$$p(w|d) = \sum_{z \in Z} p(w|z)p(z|d), \quad (15)$$

where $p(w|z)$ and $p(z|d)$ are prior probabilities. The EM algorithm is used to estimate the maximum similarity as follows:

$$L = \sum_{d \in D} \sum_{w \in W} n(d, w) \log\{p(d)p(w|d)\}, \quad (16)$$

where $n(d, w)$ is the frequency of occurrence. Through training, we get the following results:

$$\begin{aligned} p(z|w, d) &\leftarrow \theta_d(z)\phi_z(w), \\ \phi_z(w) &\leftarrow \sum_d n(d, w)p(z|w, d), \\ \theta_d(z) &\leftarrow (\alpha - 1) + \sum_w n(d, w)p(z|w, d), \end{aligned} \quad (17)$$

where α is a parameter. The corresponding maximum similarity can be written as follows:

$$L = \sum_{d \in D} \sum_{w \in W} n(d, w) \log(d, w) - \gamma \sum_{z \in Z} \sum_{w \in W} (p(z|w) - \rho(w))^2. \quad (18)$$

Based on this, the color mapping is realized, and the colors are sorted according to Figure 3 to construct the differentiation degree.

3. Experiment Result and Analysis

The experiment is composed of visible infrared hyperspectral data and software simulation data [28], as shown in Figure 4, including grassland, sand, vehicles, buildings, and other typical targets. It can be seen from the figure that the pixel values displayed by different ground objects in different bands are different, and there are also differences in the pixel values of ground objects in the same band, which is the basis of band selection. We normalized the hyperspectral images to $512 \times 512 \times 300$.

3.1. Display of Spectral Curves of Typical Ground Objects.

In order to show the spectral curves of typical ground objects, the spectral curves of leaves and sky are selected to display, as shown in Figure 5. The horizontal axis represents the band number and the vertical axis represents the pixel value. It can be seen that the same kind of features has strong similarity, and different types of features have differences. Although the internal targets have volatility, the overall volatility is small. In the areas of leaves and sky, the most significant area is concentrated in the 0–300 band. In the area of sky, the pixel value of 0–300 band reaches saturation state. Based on the above analysis, the sky and leaves can be effectively distinguished.

3.2. Band Clustering Effect. In order to verify the accuracy of moving target extraction, we introduce the OA [29] overall accuracy and kappa coefficient. OA represents the proportion of samples with correct classification to all samples. Kappa is an index used for the consistency test [30]:

$$\begin{aligned} \text{OA} &= \frac{1}{N} \sum_{i=1}^r x_{ii}, \\ \text{kappa} &= \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \times x_{+i})}. \end{aligned} \quad (19)$$

The band selection effect of real scene hyperspectral image is shown in Figure 6, and the band selection effect of simulated scene hyperspectral image is shown in Figure 7. It can be seen from the figure that the effect of the real image is slightly lower than that of the simulated image, which is due to the stable noise and spectral curve contained in the simulated image. However, the information contained in the real image is more complex and has a certain volatility.

The sparse nonnegative matrix factorization (SNMF) algorithm [6] transforms the problem of band selection into the problem of sparse decomposition, and it has a certain effect to extract significant spectral images. The fast volume gradient (FVG) algorithm [13] establishes the model according to the gradient to realize the band selection, and the effect is better for the region with obvious boundary. The variable precision neighborhood (VPN) algorithm [17] constructs fuzzy sets according to the relationship between adjacent pixels to realize band selection. The fast and late low rank (FLLR) algorithm [20] introduces the idea of low rank to calculate the redundancy between bands and realize band

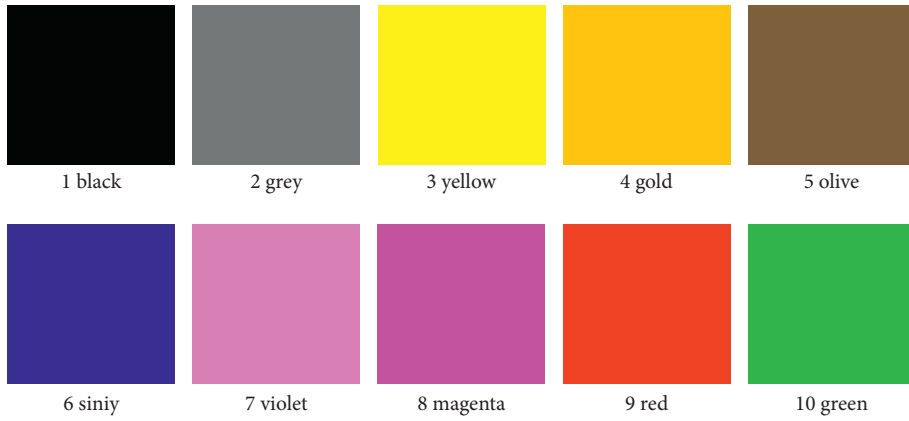


FIGURE 3: Mapping color sequence graph.

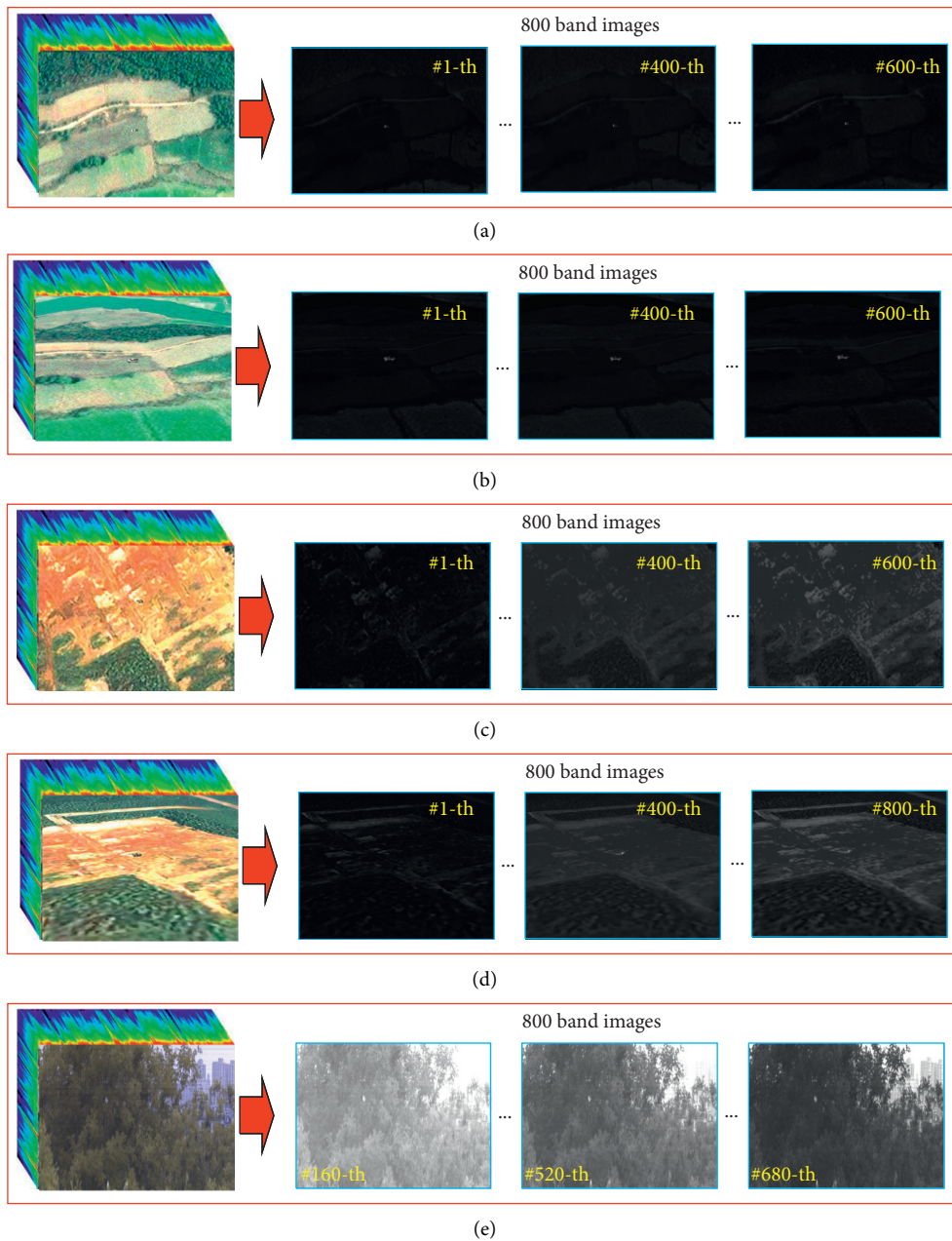


FIGURE 4: Data display: (a) simulated overlook image sequence; (b) simulated sideview image sequence; (c) simulated overlook image sequence; (d) simulated sideview image sequence; (e) real image sequence.

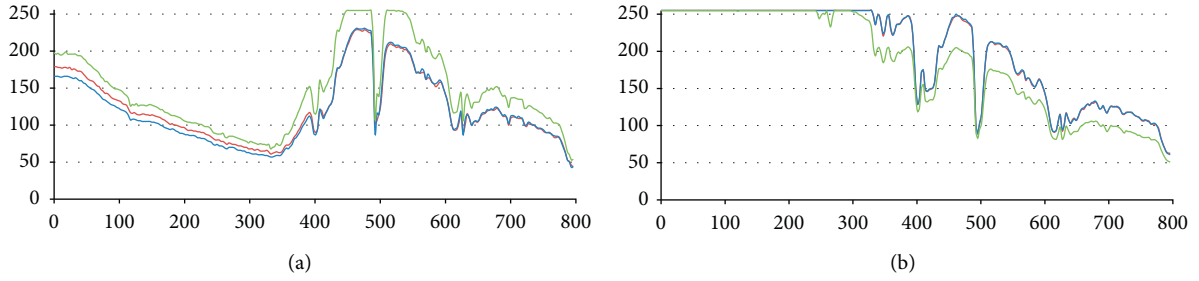


FIGURE 5: Display of spectral curves of typical ground objects: (a) spectral curve of leaves; (b) spectral curve of sky.

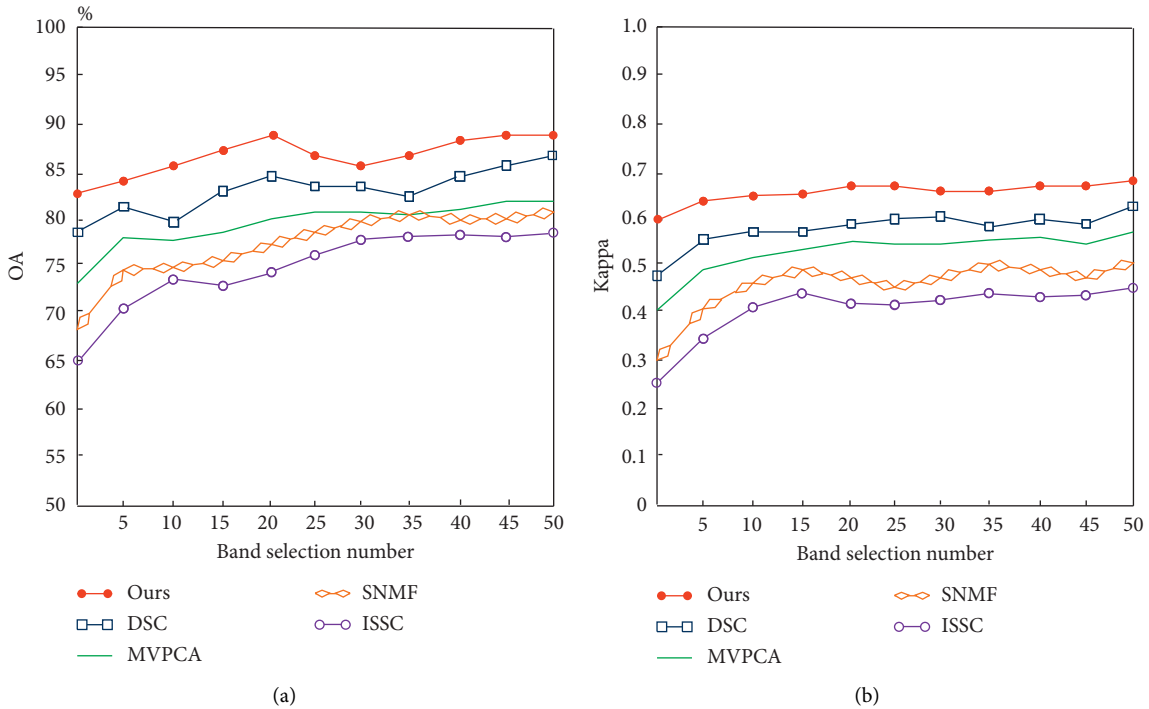


FIGURE 6: Experimental results of real hyperspectral sequences: (a) OA; (b) kappa.

selection. In this paper, a subspace clustering algorithm is proposed based on depth confrontation, which fully considers the correlation of bands and optimizes the loss function to achieve band selection. On the real data set, OA reaches 90% and kappa reaches 0.67. OA and kappa are 96% and 0.92, respectively, in the simulation data set, which reach better results.

3.3. Target Detection Effect. On the basis of the optimal band selection, different target material clustering algorithms are used for comparison. The detection results of real data and simulated data are shown in Figure 8. Hu et al. [31] proposed the SVM algorithm to extract image features for clustering to achieve enhancement. Han et al. [32] constructed CNN to extract target features. Shi and Pun [33] built multiscale RESNET to realize target detection. Li et al. [34] detected targets based on boundary features. The above algorithms analyze the target from the perspective of morphology to

achieve target detection. Based on clustering, in this paper, we construct a visual perception model to detect the target and use the difference of visual mapping to measure the detection rate of the target, which has a good effect, and the ROC curve value is the highest.

The proposed algorithm constructs the mapping model of visual perception by fusing images with three bands. The mapping results of real data using different band combinations are shown in Figure 9(a). {0, 38, 187} segment maps the leaf region to red, but it cannot distinguish the building and sky areas effectively. {1, 161, 35} spectrum segment can distinguish the building area from other areas, which verifies the effectiveness of the proposed algorithm. The mapping results of simulated data using different band combinations are shown in Figure 9(b). {0.4, 8.0, 12} spectrum can distinguish grassland from land but cannot distinguish grassland from vehicle. {0.6, 0.6, 10.4} bands can extract vehicles effectively and suppress grassland and land areas. The effectiveness of the proposed algorithm is verified.

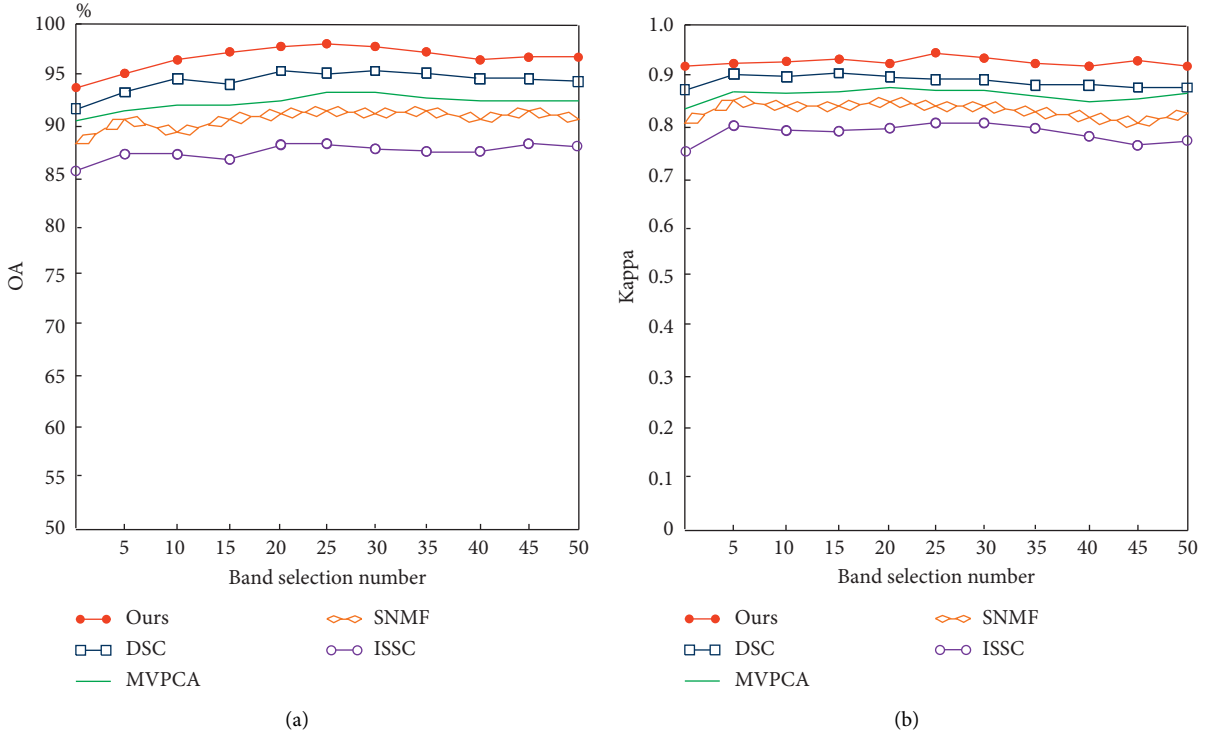


FIGURE 7: Experimental results of simulated hyperspectral sequences: (a) OA; (b) kappa.

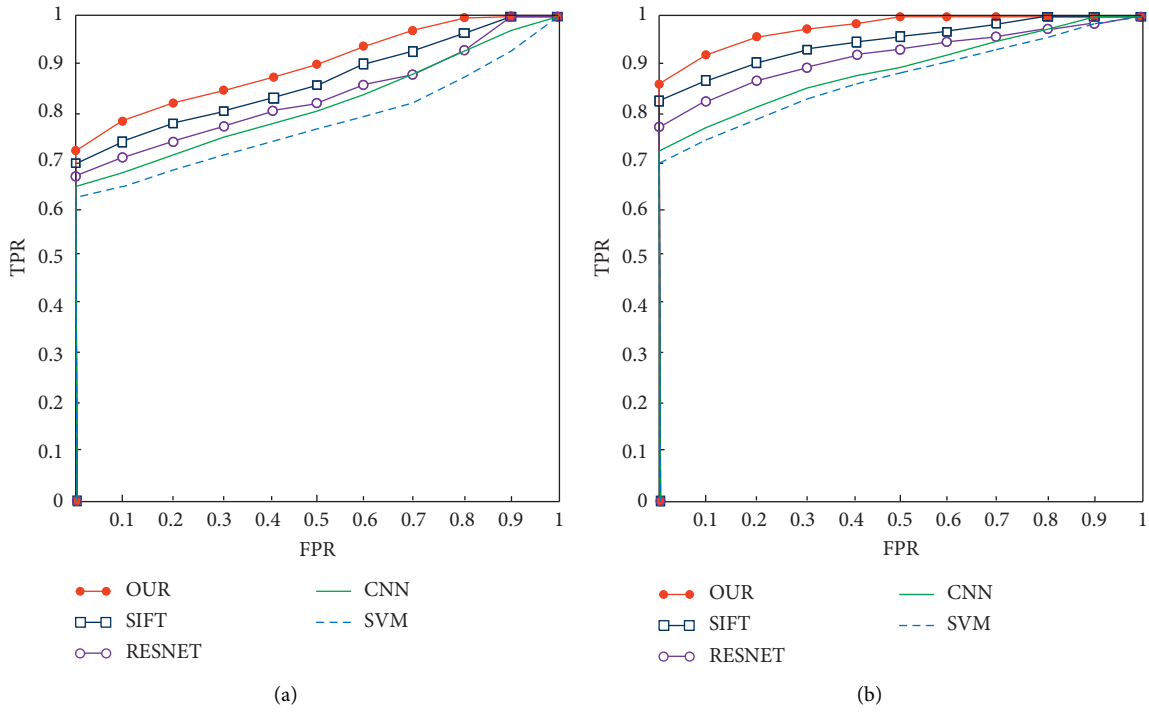


FIGURE 8: Roc curve: (a) real data; (b) simulated data.

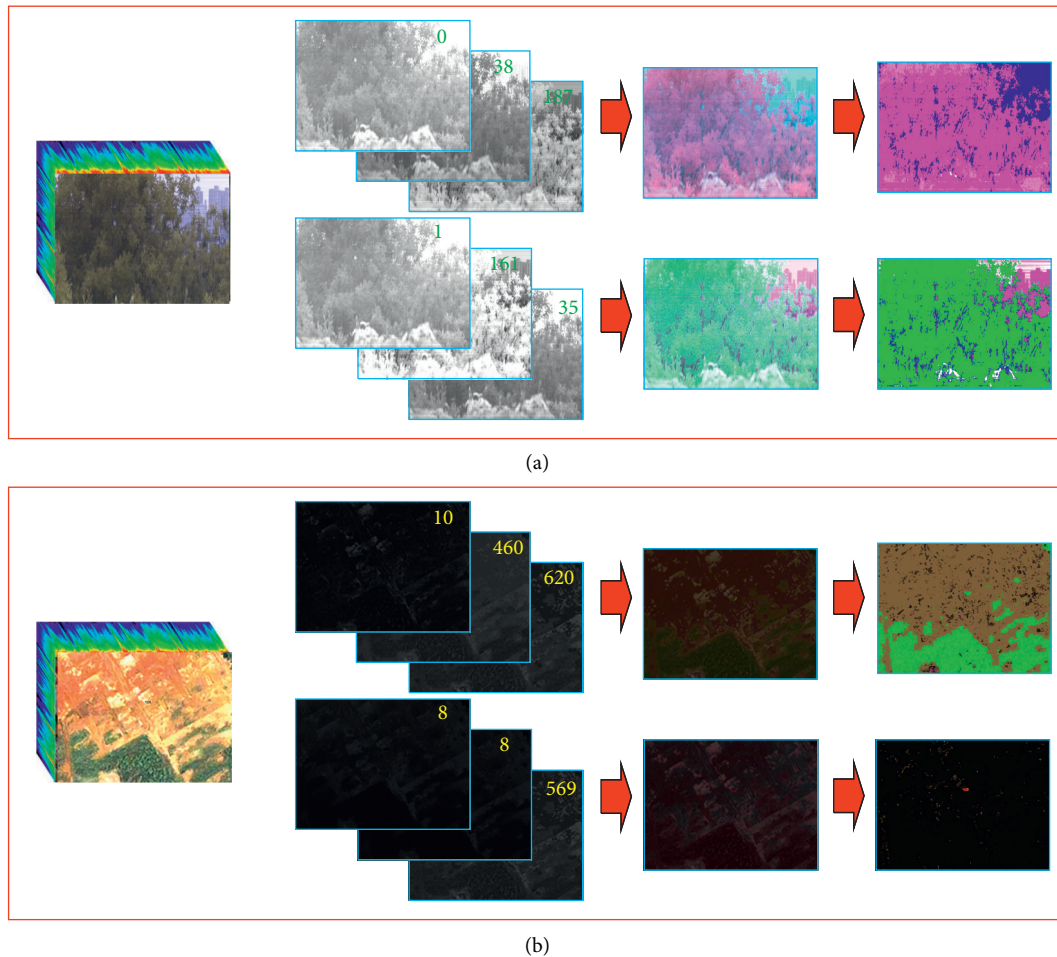


FIGURE 9: Target detection experimental results: (a) real data; (b) simulated data.

4. Conclusion

Hyperspectral images have spatial resolution and inter-spectral resolution, which plays an important role in material recognition. Aiming at the difficulty of hyperspectral band selection, a deep adversarial subspace clustering network is constructed to select the representative band, which can select a representative band. From the perspective of psychology, a color perception model is constructed to highlight the significant areas. Experiments show that the proposed algorithm has good results. On this basis, it can carry out material recognition of typical targets and hidden targets.

Data Availability

This paper experimented with two databases. One is from https://figshare.com/articles/dataset/Main_zip/2007723/3. The other is the real hyperspectral database. The data are too large and can be obtained from the corresponding author upon request. The database is for scientific research only.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Postdoctoral Science Foundation of China under grant no. 2020M682149.


References

- [1] V. K. Venugopal, K. Vaidhya, M. Murugavel et al., "Unboxing AI - radiological insights into a deep neural network for lung nodule characterization," *Academic Radiology*, vol. 27, no. 1, pp. 88–95, 2020.
- [2] S. Sawant and M. Prabukumar, "A survey of band selection techniques for hyperspectral image classification," *Journal of Spectral Imaging*, vol. 9, 2020.
- [3] C. Chang, Y. Kuo, S. Chen, C. C. Liang, K. Y. Ma, and P. F. Hu, "Self-mutual information-based band selection for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, 2020.
- [4] W. Sun, J. Peng, G. Yang, and Q. Du, "Fast and latent low-rank subspace clustering for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 6, pp. 3906–3915, 2020.
- [5] H. Yang, Q. Du, H. Su, and Y. Sheng, "An efficient method for supervised hyperspectral band selection," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 1, pp. 138–142, 2010.

- [6] W. Di, L. Zhang, D. Zhang, and Q. Pan, "Studies on hyperspectral face recognition in visible spectrum with feature band selection," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 40, no. 6, pp. 1354–1361, 2010.
- [7] J.-m. Li and Y.-t. Qian, "Clustering-based hyperspectral band selection using sparse nonnegative matrix factorization," *Journal of Zhejiang University - Science C*, vol. 12, no. 7, pp. 542–549, 2011.
- [8] F. Samadzadegan and F. Mahmoudi, "Optimum band selection in hyperspectral imagery using swarm intelligence optimization algorithms," in *Proceedings of the 2011 International Conference on Image Information Processing*, pp. 1–6, IEEE, Shimla, India, November 2011.
- [9] Q. Du, J. Bioucas-Dias, and A. Plaza, "Hyperspectral band selection using a collaborative sparse model," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pp. 3054–3057, IEEE, Munich, Germany, July 2012.
- [10] R. Hedjam and M. Cheriet, "Hyperspectral band selection based on graph clustering," in *Proceedings of the 2012 11th International Conference on Information Science, Signal Processing and Their Applications (ISSPA)*, pp. 813–817, Montreal, Canada, July 2012.
- [11] J. Feng, L. Jiao, X. Zhang, and T. Sun, "Hyperspectral band selection based on trivariate mutual information and clonal selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4092–4105, 2013.
- [12] R. Nakamura, L. Fonseca, J. Dos Santos, R. d. S. Torres, X. Yang, and J. P. Papa, "Nature-inspired framework for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2126–2137, 2013.
- [13] H. Su, Q. Du, G. Chen, and P. Du, "Optimized hyperspectral band selection using particle swarm optimization," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2659–2670, 2014.
- [14] X. Xiurui Geng, K. Kang Sun, L. Luyan Ji, and Yongchao Zhao, "A fast volume-gradient-based band selection method for hyperspectral image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 7111–7119, 2014.
- [15] S. Jia, G. Tang, J. Zhu, and Q. Li, "A novel ranking-based clustering approach for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 88–102, 2015.
- [16] S. Patra, P. Modi, and L. Bruzzone, "Hyperspectral band selection based on rough set," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 10, pp. 5495–5503, 2015.
- [17] J. Feng, L. Jiao, T. Sun, H. Liu, and X. Zhang, "Multiple kernel learning based on discriminative kernel clustering for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6516–6530, 2016.
- [18] Y. Liu, H. Xie, L. Wang, and K. Tan, "Hyperspectral band selection based on a variable precision neighborhood rough set," *Applied Optics*, vol. 55, no. 3, pp. 462–472, 2016.
- [19] X. Cao, C. Wei, J. Han, and L. Jiao, "Hyperspectral band selection using improved classification map," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 2147–2151, 2017.
- [20] V. Shah, S. Bhattacharjee, S. Silvestri et al., "Designing sustainable smart connected communities using dynamic spectrum access via band selection," in *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments*, pp. 1–10, Delft, Netherlands, November 2017.
- [21] Q. Wang, F. Zhang, and X. Li, "Optimal clustering framework for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. 5910–5922, 2018.
- [22] F. Xie, F. Li, C. Lei, and L. Ke, "Representative band selection for hyperspectral image classification," *ISPRS International Journal of Geo-Information*, vol. 7, no. 9, p. 338, 2018.
- [23] W. Sun, G. Yang, J. Peng, and Q. Du, "Hyperspectral band selection using weighted kernel regularization," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 9, pp. 3665–3676, 2019.
- [24] W. Sun, J. Peng, G. Yang, and Q. Du, "Correntropy-based sparse spectral clustering for hyperspectral band selection," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 3, pp. 484–488, 2019.
- [25] R. M. Torres, P. W. T. Yuen, C. Yuan, J. Piper, C. McCullough, and P. Godfree, "Spatial spectral band selection for enhanced hyperspectral remote sensing classification applications," *Journal of Imaging*, vol. 6, no. 9, p. 87, 2020.
- [26] W. Sun, G. Yang, J. Peng et al., "A multiscale spectral features graph fusion method for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–12, 2021.
- [27] S. Patra and B. Barman, "A novel dependency definition exploiting boundary samples in rough set theory for hyperspectral band selection," *Applied Soft Computing*, vol. 99, Article ID 106944, 2021.
- [28] Pedram Ghamisi and Stuart Phinn, "Fusion of LiDAR and Hyperspectral Data," 2015.
- [29] S. Qiu, K. Cheng, L. Cui, D. Zhou, and Q. Guo, "A moving vehicle tracking algorithm based on deep learning," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–7, 2020.
- [30] S. Qiu, Y. Tang, Y. Du, and S. Yang, "The infrared moving target extraction and fast video reconstruction algorithm," *Infrared Physics & Technology*, vol. 97, pp. 85–92, 2019.
- [31] R. Hu, X. Zhu, Y. Zhu, and J. Gan, "Robust SVM with adaptive graph learning," *World Wide Web*, vol. 23, no. 3, pp. 1945–1968, 2020.
- [32] M. Han, R. Cong, X. Li, H. Fu, and J. Lei, "Joint spatial-spectral hyperspectral image classification based on convolutional neural network," *Pattern Recognition Letters*, vol. 130, pp. 38–45, 2020.
- [33] C. Shi and C. M. Pun, "Multi-scale hierarchical recurrent neural networks for hyperspectral image classification," *Neurocomputing*, vol. 294, pp. 82–93, 2018.
- [34] Y. Li, L. Fan, and W. Xie, "TGSIFT: robust SIFT descriptor based on tensor gradient for hyperspectral images," *Chinese Journal of Electronics*, vol. 29, no. 5, pp. 916–925, 2020.

Research Article

Impact of Check-In Data on Urban Vitality in the Macao Peninsula

Chen Pan,¹ Junling Zhou,² and Xiaohua Huang³ 

¹Faculty of Innovation and Design, City University of Macau, Macau 99078, China

²Guangdong Polytechnic Normal University, Guangzhou, Guangdong Province 510665, China

³School of Design, NingboTech University, Ningbo 315100, China

Correspondence should be addressed to Xiaohua Huang; hxx@nit.zju.edu.cn

Received 7 July 2021; Revised 24 July 2021; Accepted 14 August 2021; Published 7 September 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Chen Pan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High activity is an important manifestation of the stable development of urban social economy. Quantitative research on urban development based on the geographical label perception of urban vitality is a new technical means and way to study urban vitality. In this paper, points of interest and Weibo check-in geographic markers are used to analyze urban vitality indicators and urban vitality distribution patterns. Through the application of different indexes, the ordinary linear regression and spatial autoregressive models between urban vitality and built environment are established to explore the factors that affect urban vitality. Results of the research show that interest points and social media check-in data can better indicate urban vitality. The urban vitality of the Macao Special Administrative Region is mainly affected by the density of land use, buildings, and public transportation.

1. Introduction

With the profound transformation of urban society, the importance of cities is becoming a hot topic in urban geography, urban planning, urban economy, and urban management. Jacobs believes that the connection between people's activities and places of residence, as well as the diversity of cities, will make the city important [1]. In the view of Montgomery, a lively space should be made available to a variety of crowded activities, so as to attract different groups of people.

The measurement of the importance of cities is considered the focus of related research. Most traditional studies use research methods to study the importance of cities [2]. Specifically, GPS is used to record the number of active residents in the research and measurement community. This survey method can explain the importance of the surrounding area. However, it is found to be defective because of its time-consuming and labor-consuming characteristics and the inability to cover a large area [3]. Therefore, the ordinary linear regression and spatial autoregressive models between urban vitality and built environment are established

to construct different indexes, thus revealing the factors that affect urban vitality [4].

In recent years, with the continuous integration of information technology and positioning technology, big data using geographical location labels (such as mobile signals, point of interest, and social media information) can be used to display a large amount of data about human activities, provide methods to identify physical and social locations in cities, and measure the importance of cities. Because of the advantages and rapid technological improvement, mobile signal data are often adopted to highlight the key city activities [5]. Long Ying takes mobile data as a medium of human activities and collects all signals from mobile phone stations to express vitality and establish an urban importance index system. However, mobile signal data also have some shortcomings, such as the concerns of privacy, limited access, and relatively high costs. The focus, with rich semantics, wide distribution, and easy access, has gained wide acceptance in urban research. The information involved includes services that can represent urban priorities, such as leisure, entertainment, hotels, and shopping. On the other hand, urban economic vitality can be measured by the size of

above-mentioned indexes and the small amount of catering POIs. In contrast to field data, social media feedback includes details on the location and type of events in different regions, which reflect the vitality of the city. Wu et al., in the literature, suggest the use of Weibo check-in information to reveal the potential importance of the city [6]. Overall, the categories of interest points and Weibo check-in locations can expand the criteria for measuring urban vitality. However, different geographical information may increase the vitality of the city, and the spatial patterns represented by different data can be the basis for measuring the vitality of a city. Therefore, it is necessary to deepen the analysis on its similarities and differences.

Clarifying the mechanism of urban vitality is another key point of this paper. The time and place in people's daily life are intertwined with the surrounding environment, while the living environment is a main factor affecting the vitality and importance of the city [7]. It is found that the vitality of the city is closely related to the living environment of the city, and the land use component is the basic condition of the urban activity area [8]. Different types of land planning will have different effects on the importance of cities, and mixed land use can increase the vitality of cities [9]. MEHTAV is a street full of diversities, which can be characterized by public areas, road facilities, road width, shadows, road traffic, street traffic, architectural features, individual store windows, and geographical features [10]. Therefore, the impact on road vitality cannot be ignored, and buildings are also a key to maintain urban vitality. Ye et al. refer to the method to improve urban vitality with building occupancy spatial indicators and building floor spatial indicators in the literature [5]. Therefore, a wide range of spatial data, such as land use, transport networks, and buildings [11], are used to clarify the impact mechanism of urban environmental indicators including land use, transport networks, and buildings.

With the popularity of intelligent communication devices and the continuous maturity of spatial positioning technology, location-based services (LBSs) have been widely used. The geographical location dimension builds a bridge between the real world and online social network. It aims to share the user's current location and activity information and form a social network based on location service networks (LBSN). Through the movement of the user's location, the sign-in data connect the scattered spatial nodes into a complex network with the track. It resembles a complex network built on geographic space, with transportation networks and road networks. With the popularization of network information technology, social platforms such as Facebook, Twitter, Weibo, and WeChat have produced an endless stream of spatiotemporal data, which contains complex information. This requires rigorous people to explore effective methods to mine these big data. Crowdsourced geographic data processing and mining is a current research hotspot. For example, Wang Bo constructed the public perception sentiment index by using the Weibo check-in data to analyze the spatial and temporal pattern of residents' response to floods. [12] Wang Di used different statistical analysis methods, such as kernel density estimation and

focus analysis, to analyze the characteristics of crowd activity from the perspectives of time and space in the urban Weibo attendance data [13]. There must be some close relationship for such an order of magnitude of thousands of nodes and huge components of a complex network between the edges of the nodes. In the entire network, many compact connected block networks are distinguished to divide sub-intervals, which is to divide the corresponding interactive block structure according to the attribute weight. Block structure is an essential property of complex networks. It is formed by the interaction of the elements in the organic system and the relatively stable structure of each element. These relationships include, but are not limited to, the function of the classification structure based on the interaction between the community structure and so on. By mining, we can obtain the fixed subnetwork community structure with high cohesion and low coupling in complex spatial networks. Based on the selection of microblog check-in data mining, this paper will include time attribute data and attribute data in the data and directly reflect the current tourism situation of urban residents. This paper combines a huge amount of data. As a data source of urban block structure mining, it can directly provide a basis for urban spatial structure division through users' dynamic travel. This is of great significance for guiding urban construction and evaluating existing planning schemes.

In this paper, Weibo media check-in data and point of interest (POI) data are utilized to refer to and measure urban vitality and conduct urban planning analysis [14]. Linear regression models and spatial autoregressive models are constructed using land use, urban transportation, buildings, and other built-in environmental factors [15]. It shows the environmental impact mechanism of urban vitality and the stability of the impact mechanism [16]. The study shows that the data scattered in different geographical locations have similar urban spatial patterns [17]. The city's vitality is largely affected by land use, building congestion, and the number of bus stops [18]. This study evaluates the vitality of the current city on the Macao Peninsula [19], discusses the trend of urban space development, determines the mechanism that affects the vitality of the city [20] and the importance of urban vitality to the promotion of urban development, and eventually puts forward targeted policies [21].

2. Methods

2.1. Urban Viability Test. To measure urban vitality is one of the important steps to determine the area layout. This paper will determine key indicators based on roads, which will be described by the government or regional planning department according to relevant information [22] to better study the distribution of key areas. In this paper [23], two kinds of geography markup data (POI and social media sign-in data) were used to calculate the space of two kinds of regional viability was calculated, highlight the city vigor, and analyze the space distribution pattern.

2.1.1. POI-Based Urban Viability Indicators. The point of interest (POI) is the main place where people's daily activities take place, and the POI data have good spatial scale adaptive characteristics. POI data of urban facilities contain information revealing the wide intersections between urban population and the surrounding environment. In this research, the POI data of reviews and Amed maps will be synthesized to calculate the street POI density and the vitality of the Macao Peninsula. The formula is

$$V_{\text{poi}} = \frac{P_{\text{street}}}{S_{\text{street}}}, \quad (1)$$

where P_{street} represents the number of all POIs in the region and S_{street} means the area in the region.

2.1.2. Indicators of Urban Vitality Based on Social Media Check-In. Social media check-in information represents a crowd's preference for a particular type of activity or location and can capture some ways of life of people. Here, Sina Weibo's check-in information will be used to calculate nearby social media check-in density information and show the vitality of the city. The formula is

$$V_{\text{checkin}} = \frac{C_{\text{street}}}{S_{\text{street}}}, \quad (2)$$

where C_{street} in the formula represents the number of checks in the area and S_{street} represents the area in the region.

2.1.3. Moran's I. In this paper, Moran's I is used to analyze the autocorrelation of urban importance and indicate the traditional distribution pattern of spatial importance. The formula is

$$\left\{ \begin{array}{l} I = \frac{n}{S_0} \frac{\sum_{i=0}^n \sum_{j=0}^n W_{i,j} Z_i Z_j}{\sum_{i=0}^n Z_i^2}, \\ S_0 = \sum_{i=0}^n \sum_{j=0}^n W_{i,j}. \end{array} \right. \quad (3)$$

The formula shows the deviation between the attribute of the element i and its average value and the spatial weight between the elements i and j . n is equal to the sum of elements; it is the aggregation of the weights of all spatial elements. Moran's $I > 0$ denotes spatial autocorrelation: the larger the value, the more obvious the spatial correlation; Moran's $I < 0$ denotes spatial negative correlation: the smaller the value, the more obvious the spatial dispersing; and Moran's $I = 0$ denotes that the space is random.

2.2. Built Environment. The built environment includes urban buildings and artificial sites, and artificial construction is an important factor affecting the main environment of the city. The built environment described in this paper is made up of land use, transportation, and buildings.

2.2.1. Land Use. Based on the spatial data of land use, this paper calculates the ratio of commercial land, residential land, government land, and industrial land. Shannon's entropy is used to describe the degree of mixed land use. The higher the entropy value, the higher the land utilization ratio. Similarly, the lower the entropy value, the lower the amount of relative mixture. The formula is

$$\text{Entropy} = \sum_1^n P_i \times \ln(P_i). \quad (4)$$

This formula shows the proportion of land area of a certain land use type in a specific area of each land type. n is the number of land classification applications.

2.2.2. Traffic Network. The spatial boundaries of traffic networks, such as roads and buses, have a significant impact on the importance of cities. In this paper, traffic network data are used to calculate road congestion, road network congestion, bus station congestion, and other indicators. The curvature of the road extends to the length of the entire road per unit area. Road network constraints refer to the number of routes in each region. Bus parking restrictions are expressed as the number of bus stations in each service area.

2.2.3. Buildings. The spatial composition of buildings is also an important factor useful in the definition of a city. The following parts of the study will introduce two indicators of building probability, namely, building ratio and land use right probability. Among them, the building area rate of buildings is expressed by the formula. As shown in the formula, the probability of land use right is the percentage of the area occupied by the building block.

$$\left\{ \begin{array}{l} \text{FSI}_i = \frac{F_i}{A_i}, \\ \text{GSI}_i = \frac{B_i}{A_i}, \end{array} \right. \quad (5)$$

where F_i is the floor area of all buildings on the street; B_i represents the area summary of the buildings on the street; and A_i represents the area of the block.

As shown in Table 1, the built environment is measured by 10 factors, including commercial land, industrial land, government land, land use, road network density, road network building ratio, bus station permeability, building area-to-building area ratio, building area, and soil permeability of buildings.

2.3. Linear Regression Model. First, multiple linear regression is carried out to analyze the impact of built environment on urban vitality. Linear regression (also known as quadratic regression) is one of the most widely used models in routing analysis, considering its ability to reveal the linear relationships of objects. The formula is

TABLE 1: Completion of environmental indicators.

Built environment	Indicators	Number of cases	Scope	Average value	Standard deviation	Variance
Land use	Commercial land ratio	140	0~0.841	0.404	0.172	0.414
	Residential land ratio	140	0~0.344	0.112	0.086	0.294
	Government land use ratio	140	0~0.250	0.03	0.033	0.182
	Industrial land ratio	140	0~0.353	0.082	0.08	0.283
	Other land use ratio	140	0~1.00	0.371	0.209	0.457
Traffic network	Land mixed use	140	0~1.489	0.952	0.514	0.717
	Road network density	140	0~3.171	0.763	0.361	0.601
	Network node density	140	0~99.803	27.33	13.876	3.725
	Bus site density	140	0~5.466	1.099	1.007	1.004
Buildings	Building volume	140	0~9.703	2.72	1.994	1.412
	Density of buildings	140	0~0.853	0.351	0.21	0.458
	Construction area	140				

$$y = \beta_0 + \sum_{j=0}^m \beta_j x_j + \varepsilon. \quad (6)$$

In the above formula, y is the dependent variable; x_j is the independent variable; β_j is the regression coefficient of the corresponding single variable; M represents the sum of the m of the independent variables involved in the control; β_0 is the initial initialization coefficient; and ε is the error value.

In this research, we will create a roadmap for POI and social media check-in rates, respectively. Both the dependent variables in the table are POIs (Table Model 1 in Table 2), and the density of social media check-in (Table Model 2 in Table 2) indicates urban vitality. Considering the important differences between the two cities, the continuity model is further adopted to prove the importance of establishing the environmental impact.

2.4. Spatial Autoregressive Model. Linear regression analysis does not consider spatial autocorrelation and spatial distribution of built environment. Therefore, in this paper, we will use the spatial model in practice to further clarify the impact of urban importance on the environment. The formula is

$$\begin{cases} y = \rho W_1 y + \beta_1 X + \mu, \\ \mu = \lambda W_2 \mu + \varepsilon, \\ \varepsilon: N(0, \delta^2 I). \end{cases} \quad (7)$$

In the formula, y represents the dependent variable; ρ is the $W_1 y$ coefficient of the spatial lag term; x is the independent variable; β is the regression coefficient of corresponding independent variable; μ is the error value; ε is the random error of variance for service mean value 0; W_1 and W_2 in the residual matrix table represent spatial trend and dependent variable, respectively; and λ is the coefficient of spatial error setting.

Different settings can be made according to the parameters in the above formula, and three kinds of spatial autoregressive models can be produced, including spatial lag model, spatial error model, and regression coefficient of spatial error term.

- (1) When $\rho \neq 0$ and $\lambda = 0$, the model is the spatial lag regression model. The spatial lag regression model considers the spatial correlation of dependent variables. The main point of this paper is that the importance of the neighborhood is not only affected by the dynamic elements but also by the surrounding environment. This suggests the necessity to consider the spatial correlation of vitality.
- (2) When $\rho = 0$ and $\lambda \neq 0$, the model is the space error model. The spatial error model considers the spatial correlation of the fitting error. The result of this paper shows that the block vitality is not only affected by the relevant driving factors but also by the fitting error of the surrounding block vitality.
- (3) When $\rho \neq 0$ and $\lambda \neq 0$, the model is the space Doberman model. The Doberman spatial model not only pays attention to the spatial interval of dependent variables and errors but also concentrates on the influence of spatial variation trend of independent variables. The formula is

$$\begin{cases} y = \rho W_1 y + \beta_1 X + \lambda W_2 \mu + W_3 X \beta_2 + \varepsilon, \\ \varepsilon: N(0, \delta^2 I). \end{cases} \quad (8)$$

The parameter W_3 of the equation is equal to (7) and is the spectrum table of the spatial trend of the independent variable; β_2 is the regression coefficient during this period. The choice of spatial autoregressive model depends on the application. In this paper, the spatial autocorrelation of dependent variable regions is considered with two main indexes for Moran's I test, as shown in Table 3. The results show that both spatial importance variables are in autocorrelation. Moreover, Moran's I value of the spatial regression residuals of POI density and check-in density calculated by GeoDa software is 0.000. This result represents the saliency test of the two spatial regression residuals. Therefore, this paper chooses the spatial lag regression model to analyze the relationship between the two dynamic indexes and the built environment factors. Two models are also constructed in spatial autoregressive analysis. The dependent variables are still normalized POI density (Model 3 in Table 4) and check-in density (Model 4 in Table 4). The

TABLE 2: Results of linear regression analysis.

Project	Model 1: POI density				Model 2: Social media check-in density			
	Coefficient	<i>t</i> statistics	Standard deviation	<i>p</i> value	Coefficient	<i>t</i> statistics	Standard deviation	<i>p</i> value
Constants	-0.278	-3.106***	0.090	0.002	0.066	0.629	0.104	0.531
Building volume	-0.201	-1.513	0.133	0.133	-0.405	-2.614***	0.155	0.010
Density of buildings	0.670	5.607***	0.119	≤0.001	0.260	1.867*	0.139	0.064
Road network density	0.436	2.940***	0.148	0.004	0.014	0.080	0.173	0.936
Network node density	0.023	0.197	0.115	0.844	0.240	1.797*	0.134	0.075
Bus density	0.119	1.374	0.087	0.172	-0.157	-1.552	0.101	0.123
Commercial land ratio	0.273	2.902***	0.094	0.004	0.162	1.479	0.110	0.141
Residential land ratio	0.038	0.394	0.095	0.694	0.241	2.172**	0.111	0.032
Government land use ratio	0.004	0.033	0.118	0.974	0.097	0.707	0.137	0.481
Industrial land ratio	0.113	1.286	0.088	0.201	0.069	0.681	0.102	0.497
Land mixed use	0.111	1.621	0.068	0.107	-0.059	-0.743	0.079	0.459
Travel heat	0.006	0.058	0.097	0.954	-0.114	-1.006	0.113	0.316
R^2	0.613				0.212			
Adjusted R^2	0.579				0.144			
AIC	-97.626				-54.936			
SC	-62.326				-19.637			

Note: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$ are confidence levels.

TABLE 3: Moran index test results.

Vitality indicators	Moran index	Z value	<i>p</i> value
POI density	0.5836	11.96	≤0.001
Check-in density	0.5952	12.21	0

Note: $p < 0.001$, confidence levels.

selection of variables in the spatial autoregressive model is based on the linear regression results, that is, the variables eliminated by the stepwise multiple linear regression.

3. Experiment and Result Analysis

3.1. Research Area and Experimental Data. The test site of this article is located on the Macao Peninsula. The Macao Peninsula (Portuguese: Macao, English: Macao Peninsula) is one of the three parts that make up Macao, China. It is the main residential area of Macao and the oldest development zone in Macao, with a history of more than 400 years. Macao's historical center is located in the Macao Peninsula, which was listed as the World Cultural Heritage in 2005. The Macao Peninsula was originally an island, but as sediments had accumulated in the upper reaches of the west river, there slowly formed a sandbar facing the north and south, connecting the island to the land to create a peninsula. The Macao Peninsula is composed of five districts: Fathi Ma Tong District, San Andoni Tong District, Lobby District, Wangde Tong District, and Fengshun Tong District (Figure 1). Most of the flat land on the Macao Peninsula is obtained by landfill. The Macao Peninsula covers an area of about 9.1 square kilometers and has a population of about 388,000.

The Macao Peninsula area is about 9.1 square kilometers, and the population is about 388,000.

According to the Statistics Bureau of Macao in 2020, Macao has a total population of 682,500. Among them, most

of the Chinese were immigrants from the mainland, accounting for 97%. About 170,000 nonlocal employees are employed in Macao. Macao is a tourist city. Most of the visitors to Macao are mainlanders. Since Macao is dominated by mainlanders and the main social media application in China is Weibo, more than 80% of Macao people use Weibo. Therefore, the use of Weibo attendance data as social media data in this paper to measure the urban vitality of Macao has a certain universality.

Here, we use geo-tagged data covering the Macao Peninsula, including POI data and social media check-in information. POI data can be obtained from Amap and Dianping data. By the deadline 2019, altogether 40,654 pieces of POI data were collected. POI data include restaurants, businesses, residents, services, schools, and places of entertainment. Each piece of POI data contains information such as latitude, longitude, name, address, and category. Social media login information is captured by API. The valid date was from January 2018 to December 2019. Weibo check-in data have been preprocessed. 145,465 pieces of Weibo check-in data were received. There was a year of difference between POI data and social media, until the return, followed by the entry into a stable period of development. The data on land use and buildings are derived from the Macao Mapping and Cadastral Bureau. Road network data are derived from OpenStreetMap (OSM). The Macao Peninsula experimental area is divided into 140 blocks.

3.2. Dynamic Distribution Model of Macao Peninsula. The spatial distribution patterns of vitality on the Macao Peninsula based on POI data and social media check-in data are shown in Figures 1(b) and 2(b), respectively. Figures 1(a) and 1(b) respectively shows that the urban spatial vitality indicated by the POI data has a high activity value

TABLE 4: Spatial autoregressive results.

Project	Model 1: POI density				Model 2: Social media check-in density			
	Coefficient	Z value	p value	Standard deviation	Coefficient	Z value	p value	Standard deviation
Spatial lag items	0.559	0.061	9.107	0.000	0.679	0.067	10.079	0.000
Constants	-0.185	0.070	-2.639	0.008	-0.024	0.076	-0.319	0.749
Building volume	-0.027	0.102	-0.268	0.789	-0.158	0.115	-1.382	0.167
Density of buildings	0.444	0.094	4.720	0.000	0.040	0.102	0.390	0.696
Road network density	0.419	0.114	3.662	0.000	-0.159	0.126	-1.256	0.209
Network node density	-0.098	0.088	-1.112	0.266	0.143	0.098	1.469	0.142
Bus density	0.059	0.067	0.890	0.374	-0.128	0.074	-1.730	0.084
Commercial land ratio	0.148	0.073	2.041	0.041	0.169	0.080	2.103	0.035
Residential land ratio	-0.015	0.073	-0.205	0.837	0.189	0.082	2.298	0.022
Government land use ratio	-0.025	0.090	-0.280	0.779	0.111	0.100	1.104	0.269
Industrial land ratio	0.093	0.067	1.384	0.166	0.081	0.075	1.082	0.279
Land mixed use	0.085	0.052	1.626	0.104	-0.040	0.058	-0.696	0.486
Travel heat	-0.093	0.076	-1.214	0.225	-0.047	0.083	-0.569	0.569
R ²	0.751				0.540			
AIC	-147.734				-112.493			
SC	-109.493				-74.251			

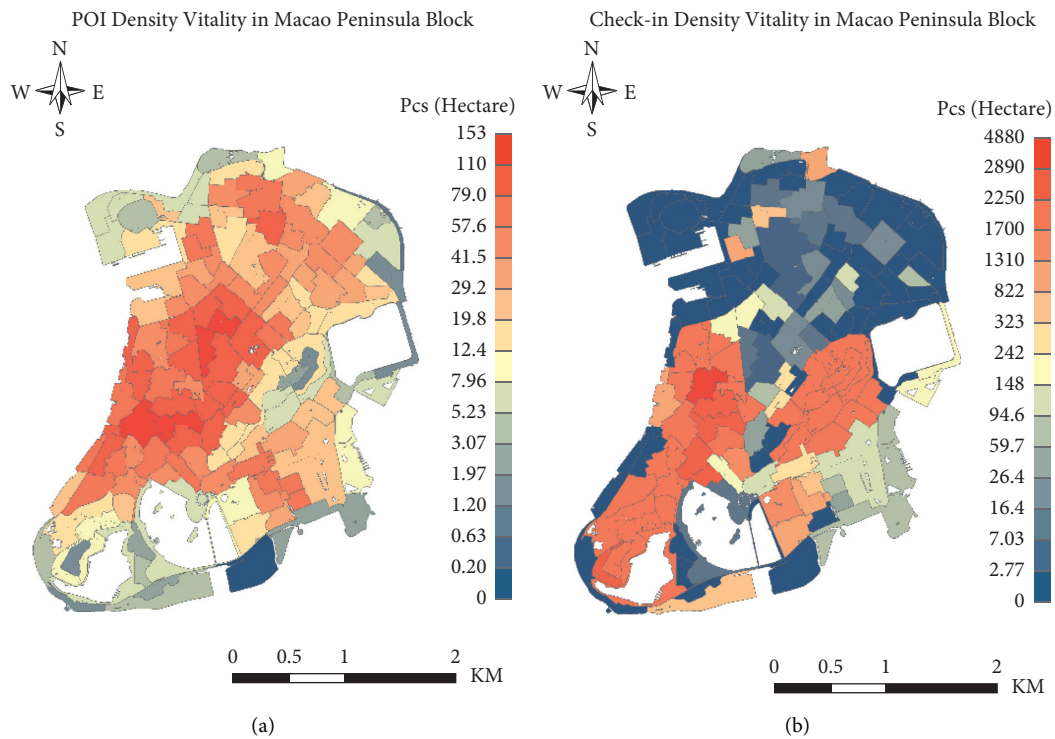


FIGURE 1: Map of POI density patterns in Macao Peninsula block: (a) POI density vitality and (b) check-in density vitality.

aggregation polar core in the mid-west part with a high-high distribution pattern. It suggests that high values surround each other; in the north and southeast parts, the high value is surrounded by the low value. Figures 2(a) and 2(b) show the urban dynamism represented by social media sign-in data. The results show that there are high active value aggregation centers in two historical areas of Macao, which are distributed in the western part and the southwest and the eastern part of the middle Macao Peninsula. The Weibo check-in density pattern map further shows that the urban

vitality represented by the sign-in density in the historical city of Macao is highly aggregated. It follows the high-high distribution pattern. 0 represents blocks with less clustering, while 1, 2, and 3 represent groups with confidence levels higher than 0.1, 0.04, and 0.01, respectively.

The test results show that an obvious imbalance exists in the spatial distribution of urban dynamic areas displayed by social media check-in information and POI data. In other words, high-energy blocks appear in historic blocks, while other areas are depressed. The POI data reveal that its vitality

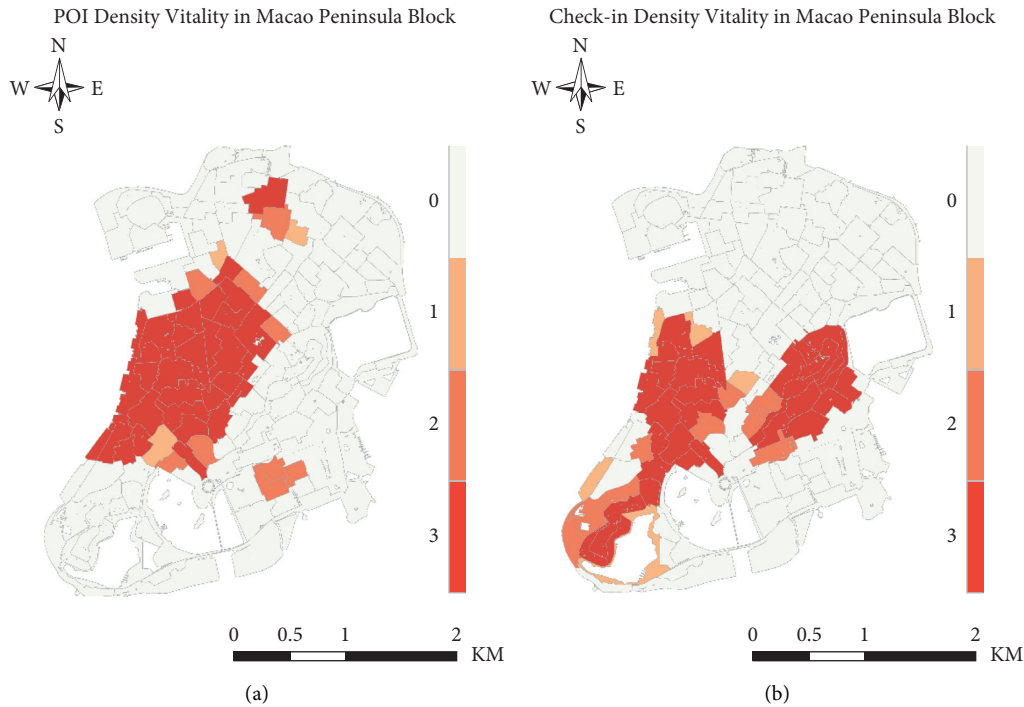


FIGURE 2: Map of Weibo sign-in density pattern in Macao Peninsula block: (a) POI density vitality pattern distribution and (b) sign-in density vitality pattern distribution.

is distributed “continuously” in space. It extends from the central and western regions to the north and southeast of the subcenter, and the vitality value of the surrounding area showed a downward trend. The two vitalities also exhibit differences in spatial aggregation. This phenomenon occurs in Macao’s historic district, where POI density and vibrant spaces are clustered in the south. It forms a balanced pattern of aggregation.

Furthermore, Figure 1 shows the Moran vitality index of the two cities. Moran’s I values of urban importance are 0.574 and 0.586, respectively, according to POI and social media check-in data. The dynamic regions of the two regions have spatial autocorrelation, both of which are expressed as the dynamic aggregation of the urban dynamic distribution regions. The test of Z and p values shows that the dynamic autocorrelation characteristics of the two urban significance indicators are statistically significant. Therefore, it is suggested that the spatial model of independent analysis of vitality can be used to analyze the specific impact of built environment on urban vitality.

3.3. Linear Regression Analysis. In this paper, we first consider the relationship between independent variables and eliminate the high variance building area of independent variables by collinearity analysis. The final variables classified are the percentage of commercial land, residential land, government land, industrial land, mixed land use, road use, road network use, bus station use, travel heat, building volume ratio, and building occupation density.

The analysis refers to the results of bit multivariate linear geometry in Table 2, which presents parameter coefficients,

test t statistics, and R^2 . The two established environmental factors that have the greatest impact on urban vitality are the appearance of buildings and the proportion of commercial land. The results of R^2 adjustment show that the interpretation rate of urban dynamic environment is 56.8% and 16.7%, respectively, based on POI and check-in data. As shown by the results of linear regression analysis, the urban vitality of architectural environment interpretation is relatively low.

The observation and experiment demonstrate that the effect of built environment on the vitality of two kinds of cities has common characteristics. Among them, there is an environmental factor with significant effects, that is, the density of building occupation, and the influence coefficient is positive, indicating that this factor generates a positive effect on urban vitality. The influence coefficient is positive, suggesting that this factor has a positive effect on urban vitality. In the future urban construction, increasing the density of buildings and commercial land ratio can enhance the vitality of the city. Among them, the partial correlation coefficient of buildings is the largest, while the buildings with high urban speed have the greatest influence on the importance of promoting the city. According to different data, the impact of the created environment on urban importance varies. For example, the popularity of POI-based public transport stations has significantly increased the importance of cities and has a significant impact on urban vitality. However, the urban vitality based on check-in data has a significant inhibitory effect. In addition, the social activities with high consumption, such as entertainment, shopping, sports, and tourism, are less affected by the density of public transport websites.

3.4. Spatial Autoregressive Analysis. The influencing factors of urban vitality are mainly obtained by performing stepwise linear regression. Table 4 shows the analysis results of the spatial lag regression model. The R^2 modes of the two spatial lag modes are 74.2% and 53.4%, respectively, proving the effectiveness of spatial lag control mode. Compared with the linear model, the interpretation rate of the model increased by 18.3% and 38.9%, respectively. The results of the lag analysis show that the smaller the values of AIC and SC are, the better the fitting effect of the regression model is. According to linear regression comparison results, the values of AIC and SC decreased by 52.6%, 74.9%, 114.6%, and 287.3%. It shows that the spatial lag regression model is superior to the linear regression model. Therefore, the spatial lag regression model can better reveal the influence mechanism of abbreviated environment on urban vitality.

The spatial lag regression coefficient expresses the degree of urban vitality affected by its spatial dependence. Among model 3 and model 4, the coefficients of spatial lag regression terms are 0.559 and 0.679, respectively. Besides, the significant level is $p \leq 0.001$, which indicates that the regional vitality is greatly affected by the vitality of adjacent regions and all remains positive. On the basis of the sign-in data, the spatial lag regression coefficient of urban vitality is greater than the POI, which indicates that its spatial autocorrelation exerts a great influence on the urban vitality under the sign-in density indication.

There is no change in the influence of building volume ratio, building occupation density, bus site density, commercial land ratio, industrial land ratio, and mixed land use. It demonstrates that the positive and negative effects of some built-up environmental factors on the importance of cities are frequent. The land density of buildings, the proportion of commercial land, and the ratio of industrial land are always positively related to vitality. Meanwhile, the higher building volume ratio has always had a negative impact on urban vitality.

Compared with the multivariate linear regression model, the influence coefficient of building area rate of the spatial lag regression model is lowered by 0.168 and 0.256, respectively. Considering the function of domain vitality, the negative effect of building volume ratio is weakened. Similarly, the effect of building occupation density on the two vitalities decreased by 0.226 and 0.220, respectively. At the same time, the effect of mixed land use on the two vitalities decreased by 0.026 and 0.074, respectively. As a result, regardless of the dynamic space autocorrelation, the building volume ratio, the density of the building, and the effect of land use on the vitality will be magnified.

4. Conclusion and Discussion

Uneven urban development can lead to problems such as unbalanced distribution of facilities and population dispersion, which can contribute to empty cities and heavy traffic. Therefore, it is of great importance to understand the current urban development situation caused by the spatial development of unbalanced areas and to deal with the urban vitality. From the perspective of urban environmental

awareness, the present study employs viewpoint (POI) and social media check-in data to create two key indicators and 12 constructive environmental indicators. Furthermore, we adopt the linear regression model (OLS) and spatial autoregressive model to analyze the characteristics of the influence of built environment on urban vitality. In the study of POI density, this phenomenon occurs in the central and western part of the Macao Peninsula with one main and two pairs of patterns. In the activity of check-in density, the experiment is summarized as follows: The urban vitality represented by various data is similar and different in spatial distribution. The vitality of the main areas of the Macao Peninsula is distributed in the central and western regions, while it is weak in the north of, southeast of, and around the island. As shown in the social media records, there is obvious "chaos" in the dynamic distribution of major cities. Besides, the vitality revealed by the POI is "continuous" in space. There also exist major differences between these two types of spatial aggregation. In the POI density space, the phenomenon occurs in the central and western part of the Macao Peninsula with one main and two pairs. In addition, the phenomenon occurs in the historical urban area of Macao in the sign-in density vitality.

The Macao Peninsula is mainly subject to land use and buildings. Among them, commercial land, building area, and bus station density generate significant effects on vitality. The proportion of mixed use of land for construction land has a negative impact. Our goal aims to strengthen the importance of industrial development zones such as Huadi Matang District and its lobby area. Attention should be paid to increasing commercial land, enhancing road network, and improving the visibility of bus stations. In the field of business and traffic, such as the lobby area of the Macao Shopping Centre, its vitality can reduce the proportion of the industrial soil and soil utilization to enhance its vitality. The construction process of some environmental factors affects only part of the key indicators. The intensity of business and road network has a positive effect on POI intensity but had no positive effect on social media audit intensity. In addition, urban planning can adopt different improvement strategies for different aspects of urban vitality.

This study helps to understand the background of urban vitality and formulate policies to improve it. Using POI and social media data to determine the importance of cities can illuminate the relationship between urban importance and many socioeconomic factors including urban dynamics. At the same time, the city's size and utility of focus have a remarkable effect on the spatial distribution of key urban areas. Finally, it is concluded that dense vitality is an important part of spatial analysis and it is necessary to study the problem of plastic space units in order to understand urban vitality.

At present, this paper only uses POI and social media data to determine the importance of cities and clarify the city's importance and the relationship between many social and economic factors, which cover vitality of the city. At the same time, the size and practicality of the city have a great impact on the spatial distribution of the key urban areas. The research needs to be extended to other cities in order to

study the uniformity of urban vitality distribution and its influencing mechanism. Finally, intensive activity is an important part of spatial analysis. It is necessary to explore plastic spatial units while understanding urban vitality.

Data Availability

The research data used to support the findings of this study are available from the corresponding author only under license and so cannot be made freely available. Requests for access to these data should be made to the author.

Conflicts of Interest


There are no potential conflicts of interest in this paper, and all the authors have seen and approved the manuscript to be published.

References

- [1] J. Jacobs, *The Death and Life of American cities*, Vintage Books, New York, NY, USA, 1961.
- [2] P. R. Mass, *Towards a Theory of Urban Vitality*, Vancouver: University of British Columbia, Vancouver, Canada, 1984.
- [3] Y. Liu, X. Liu, and S. Gao, "Social sensing: a new approach to understanding our socioeconomic environments," *Annals of the Association of American Geographers*, vol. 105, no. 3, pp. 512–530, 2015.
- [4] W. Tu, J. Cao, and Y. Yang, "Coupling mobile phone and social media data: a new approach to understanding urban functions and diurnal patterns," *International Journal of Geographical Information Science*, vol. 31, no. 12, pp. 2331–2358, 2017.
- [5] Y. Ye, D. Li, and X. Liu, "How block density and typology affect urban vitality: an exploratory analysis in Shenzhen, China," *Urban Geography*, vol. 39, no. 4, pp. 631–652, 2018.
- [6] C. Wu, X. Ye, F. Ren, and Q. Du, "Check in behaviour and spatio-temporal vibrancy: an exploratory analysis in shengzhen, china," *Cities*, vol. 77, pp. 104–116, 2018.
- [7] Y. Yang, Z. Yan, A. Yeh, J. Y. Xie, C. L. Ma, and Q. Q. Li, "Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy," *International Journal of Geographical Information Science*, vol. 31, no. 4, pp. 658–675, 2017.
- [8] W. Tu, R. Cao, Y. Yang et al., "Spatial varital variations in urban public ridership derived from GPS trajectories and smart card data," *Journal of Transport Geography*, vol. 69, pp. 45–57, 2018.
- [9] J. Lu and F. J. Zhang, "Coordinated development of historical and cultural heritage and urban vitality," *New Building*, no. 1, pp. 32–36, 2016.
- [10] Y. Ye and Y. Zhuang, "The hypothesis of spatial pattern evolution and urban vitality: a comprehensive analysis based on road, density and architectural form and functional mixing," *International Urban Planning*, vol. 32, no. 2, pp. 43–49, 2017.
- [11] L. Tang, L. Niu, Y. Xue et al., "Identification and structure extraction of urban road intersection using trajectory big data," *Journal of surveying and Mapping*, vol. 46, no. 6, pp. 770–779, 2017.
- [12] B. Wang, Z. Feng, and H. H. Sun, "Temporal and spatial analysis of urban residents' response to rainstorm and flood based on social media check-in data," *Geographical Science*, vol. 40, no. 9, pp. 1543–1552, 2020.
- [13] D. Wang and X. Zuo, "Analysis of spatial-temporal characteristics of crowd activity based on urban microblog check-in data," *Geographical Information World*, vol. 27, no. 2, pp. 28–33, 2020.
- [14] J. Shan, K. Qin, C. Huang et al., "Discussion on processing and analysis methods of Zhongyuan geographic data," *Journal of Wuhan University (Information Science Edition)*, no. 4, pp. 390–396, 2019.
- [15] L. Kong, L. Guan, Y. Ding et al., "Urban regional assessment and its key technologies," *Supported by spatial big data Bulletin of Mapping*, vol. 8, pp. 100–105, 2017.
- [16] D. Ying and Y. Zhou, "Architecture," *Urban Theory*, vol. 15, pp. 52–62, 2016.
- [17] Q. Li, *From Gematics to Urban Info Edition*, vol. 42, no. 1, pp. 1–6, 2017.
- [18] M. Lou, H. Zhang, X. Lei et al., "DBH-tree height model of natural Mongolian broad-leaved mixed forest based on spatial autocorrelation J]," *Tree Height Forestry Science*, vol. 53, no. 6, pp. 512–530, 2017.
- [19] C. Yan, J. Minhe, and T. Song, "The influence of Shanghai City Street on consumption vigor: an Analysis based on POI Diversity and intensity," *Journal of Suzhou Institute of Science and Technology (Natural Science Edition)*, vol. 13, no. 2, pp. 73–80, 2019.
- [20] D. Wang, W. Zhong, and D. Xie, "Evaluation and analysis of the built environment in shanghai, such as xie dongcan, Wang de, zhong weijing - exploration of mobile phone signaling data," *New Technology Applications*, vol. 42, no. 10, pp. 97–108, 2018.
- [21] S. Wang, Y. Liu, Z. Chen et al., "Methods of urban site scope perception under Dianping data," *Journal of Surveying and Mapping*, vol. 47, no. 8, pp. 1105–1113, 2018.
- [22] Q. Hu, M. Wang, and Q. Li, "Exploring urban hot spots and business circles by using location check-in data," *Journal of Mapping*, vol. 43, no. 3, pp. 314–321, 2014.
- [23] L. Xiang, Z. Chen, J. Wu et al., "Discussion on urban resident population grid method based on optical data and spatial adaptation model," *Journal of Geographic Information System*, vol. 19, no. 10, pp. 1298–1305, 2017.

Research Article

Pedestrian Behavior Recognition Based on Improved Dual-stream Network with Differential Feature in Surveillance Video

Yonghong Tan , Xuebin Zhou, Aiwu Chen, and Songqing Zhou

School of Intelligent Manufacturing, Hunan University of Science and Engineering, Yongzhou 425199, Hunan, China

Correspondence should be addressed to Yonghong Tan; tyh2977@huse.edu.cn

Received 16 July 2021; Accepted 11 August 2021; Published 25 August 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Yonghong Tan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the pedestrian behavior recognition accuracy of video sequences in complex background, an improved spatial-temporal two-stream network is proposed in this paper. Firstly, the deep differential network is used to replace the temporal-stream network so as to improve the representation ability and extraction efficiency of spatiotemporal features. Then, the improved Softmax loss function based on decision-making level feature fusion mechanism is used to train the model, which can retain the spatiotemporal characteristics of images between different network frames to a greater extent and reflect the action category of pedestrians more realistically. Simulation results show that the proposed improved network achieves 87% recognition accuracy on the self-built infrared dataset, and the computational efficiency is improved by 15.1%.

1. Introduction

Pedestrian action recognition is an important research direction in the field of computer vision, which has important research significance and application value in the military and civil fields such as video surveillance, intelligent transportation, motion analysis, navigation, and guidance [1–4]. Due to the poor image quality of low-cost cameras and the lack of stable and obvious feature information, the difficulty of pedestrian detection and behavior recognition is increased [5].

In order to improve the effect of pedestrian action recognition, scholars at home and abroad have also proposed many action recognition algorithms [6–12]. Li et al. proposed an action recognition method based on sparse coding and spatial pyramid feature extraction [6]. Fernando et al. proposed action recognition based on dual-tree complex wavelet transform, where support vector machine (SVM) was adopted to classify and recognize the samples' wavelet entropy [7]. In order to make full use of the complementary features in different modes, Varol et al. proposed an action recognition model based on multimodal feature fusion, which improved the recognition performance of low contrast object [8]. With the development of

hardware technology in recent years, deep learning has been widely used in the field of image processing [9–12]. At present, pedestrian detection and action recognition algorithms are usually based on deep learning network, mainly using three-dimensional convolution network, long-term and short-term memory network (LSTM), and dual-flow network to learn high-dimensional spatiotemporal features and automatically classify and recognize [10].

Kuehne et al. [11] designed an action recognition algorithm based on convolutional neural network to meet the needs of assisted driving. Ioffe and Szegedy [12] proposed an action recognition network based on a multilevel segmentation model, which improved the detection accuracy of pedestrian actions in complex backgrounds by extracting deep features from suspected regions. Pedestrian action recognition based on deep networks mostly analyzes the detected pedestrians to realize the recognition of different simple actions, such as standing, walking, squatting, and running. However, the human body action is a sequence action, and only the introduction of temporal-domain features can help improve the accuracy of recognition. Wang et al. [13] extended the original two-dimensional convolution kernel to a three-dimensional convolution kernel and proposed an abnormal behavior model based on three-

dimensional convolution, but such methods have complex parameter settings and a huge amount of parameters. LSTM uses convolutional networks to extract pedestrian features frame by frame and makes full use of the spatial-temporal characteristics of pedestrians to improve the characterization ability of behavioral actions. However, its multiscale and high-dimensional processing mode restricts the network operation speed.

As we all know, the visual cortex is mainly responsible for processing visual information in the cerebral cortex. It has two information output channels, dorsal stream and ventral stream, which correspond to spatial pathways and content pathways, respectively [14]. Inspired by this, Simonyan and Zisserman [15] creatively proposed action recognition based on dual-stream convolutional networks. The dual-stream convolutional neural network is a model that combines the spatial information network and the temporal information network. It uses the optical flow as the network input to compensate for the temporal-dimension information that the spatial network cannot capture and merges the results obtained from different models. The recognition accuracy of pedestrian behavior is improved, but the extraction process of optical flow takes a long time, which does not meet the real-time requirements of engineering development.

In order to improve the accuracy and efficiency of pedestrian detection and action recognition in video sequences under complex backgrounds, this paper proposes a fast and effective action recognition model on the basis of dual-stream convolutional networks. A comparative experiment was carried out to verify the practicability and effectiveness of the proposed algorithm. The contribution of this work can be summarized as follows:

- (1) To improve the pedestrian behavior recognition accuracy of video sequences in complex background, an improved spatial-temporal two-stream network is proposed in this study
- (2) The deep differential network is used to replace the temporal-stream network so as to improve the representation ability and extraction efficiency of spatiotemporal features
- (3) The improved Softmax loss function based on decision-making level feature fusion mechanism is used to train the model, which can retain the spatiotemporal characteristics of images between different network frames to a greater extent and reflect the action category of pedestrians more realistically

2. Related Works

2.1. Dual-Stream Convolutional Networks. The dual-stream network structure is composed of two independent spatial stream networks and temporal-stream networks, which are used to learn the spatial position information between video frames and the temporal motion characteristics in optical-flow data, as shown in Figure 1. The two networks have the same structure; each structure is composed of 3 pooling layers and 3 convolutional layers, and a nonlinear layer is

added after each convolutional layer. Although the two independent networks have the same structure, they play different roles in the dual-stream network. The input of the spatial stream network is the original image sequence, while the input of the temporal-stream network is the optical flow between adjacent data. In order to better characterize the temporal and spatial characteristics of video sequences, the dual-stream network structure designs two fusion layers, whose purpose is to fuse spatial and motion characteristics at spatial positions so that the channel responses at the same pixel positions are consistent.

The output result of the dual-stream network structure adopts cascade fusion. It is assumed that the output features of the two networks are represented as $x^A \in R^{H \times W \times D}$ and $x^B \in R^{H \times W \times D}$, respectively, where H , D , and W are the height, number of channels, and width of the feature map. The fusion operation stacks the two feature maps on the same spatial position (i, j) of the entire feature channel d , so $y_{i,j,d} = x_{i,j,d}^A$ and $y_{i,j,2d} = x_{i,j,d}^B$ can be obtained. Dual-stream network greatly improves the accuracy of behavior recognition, but it also has certain limitations. The temporal characteristic of the dual-stream network exists in the optical-flow map between adjacent frames, and the utilization information of the time dimension is limited, and the optical flow calculation complexity is relatively high. The dual-stream network cannot model the pixel-level relationship of spatiotemporal features.

2.2. Long Short-Term Memory (LSTM). LSTM network is a special recurrent neural network, which can solve the problem of long-term dependence and gradient vanishing problem of recurrent neural network [16]. The LSTM network adopts the “gate structure” to transfer the information processed at the current moment to the next moment, which can fully mine the effective information contained in the massive data. Its network structure is shown in Figure 2. The so-called gate structure refers to a sigmoid network layer (σ) and a bitwise multiplication operation.

As we all know, historical data that have occurred help to predict the probability of occurrence of the next moment. Most current recursive networks use the state of the last frame for feature representation, which obviously loses most of the dynamic information. Compared with extracting feature information from local frames, the overall features of the entire sequence can better present a global representation.

3. An Improved Dual-Stream Network for Behavior Recognition

3.1. Spatiotemporal Feature Propagation. As we all know, the dual-stream network structure is composed of two independent spatial stream networks and temporal-stream networks, respectively, inputting image sequences and optical-flow map. The calculation of optical flow is complicated and requires a lot of hardware resources, resulting in poor real-time engineering application. When the camera is fixed, the optical flow in the foreground is zero. In other words, the

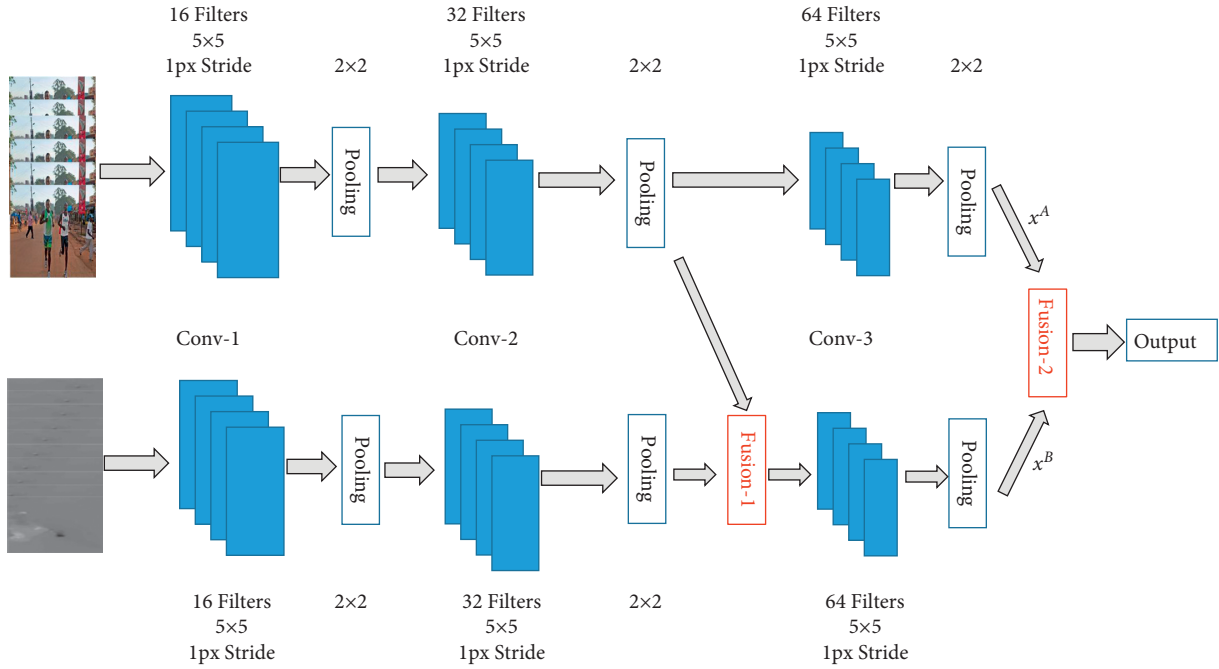


FIGURE 1: Two-stream network.

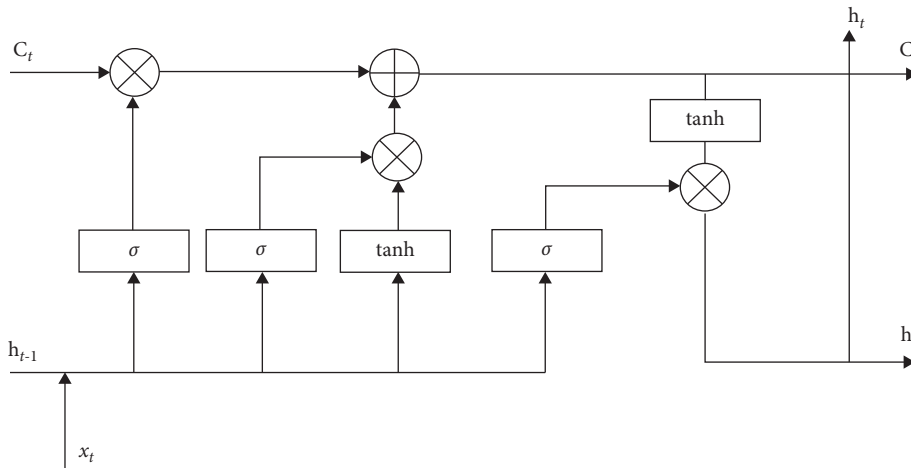


FIGURE 2: LSTM structure.

difference between images is similar to the result of optical flow. Therefore, this paper proposes a dual-stream network model based on deep differential, which uses deep differential network instead of temporal stream to obtain the interframe relationship and temporal relationship in the sequence. Deep differential is a network structure based on deep feature propagation [12]. The key frame-based feature propagation differential map can be used to replace the optical flow map as the input of the temporal-stream network, which can reduce the computational complexity and enhance the posture expression and category recognition capabilities of the feature propagation map for human actions. The improved dual-stream network structure is shown in Figure 3, where the size of the convolution kernel is 7×7 and 3×3 , respectively.

The adjacent sequences' output by the camera has a high degree of similarity, and the optical flow characteristics obtained by it are very weak. In other words, the optical flow characteristics obtained by a large number of frame-by-frame optical flow calculations are not obvious. The differential key frame proposed in [17] can quickly obtain the difference between images and improve the performance of image compression. Differential key frames include the temporal relationship between adjacent frames in the video and have similar performance to optical flow map, but they have the advantages of fast generation speed and small computational operations. Due to the problems of large redundancy and high complexity between frames, our improved dual-stream network for behavior recognition first extracts frames

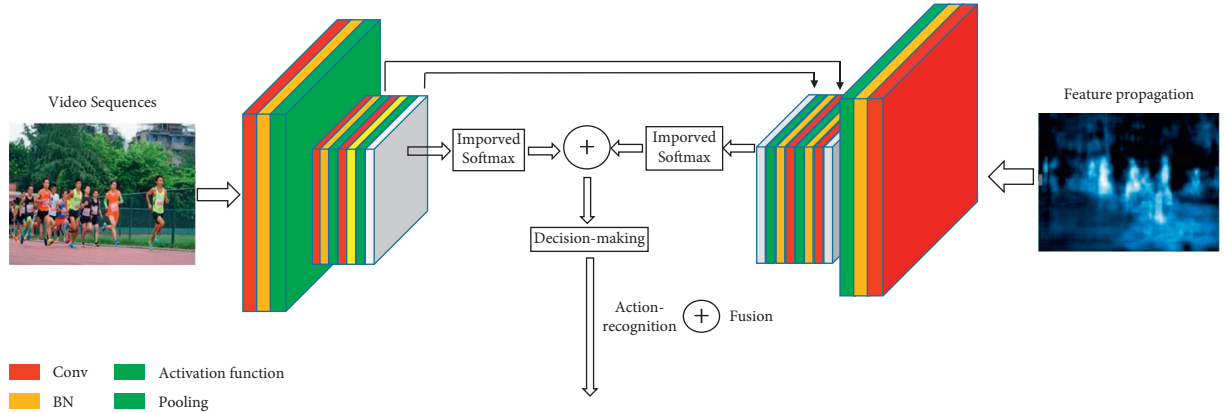


FIGURE 3: Improved two-stream network.

based on multiple time scales and uses differential feature propagation to obtain key frames of the sequence.

Assuming that an input video composed of a sequence of t frames is recorded as \mathbf{X} , each segment is first divided into T parts with the same time, and then, the key frame x_i is extracted from each part, so the entire video is recorded as $\mathbf{X} = \{x_1, x_2, \dots, x_t\}$. These key frames can be converted into differential key frames by calculating the difference between adjacent frames and is denoted as $\mathbf{Y} = \{y_1, y_2, \dots, y_t\}$; finally, the key frame and the differential key frame are input to the differential convolution network, respectively, to obtain the corresponding high-dimensional spatiotemporal feature vector $\{S_1, S_2, \dots, S_t\}$, where $S_t \in R^d$, $i = 1, 2, \dots, T$, and d is the feature dimension of the key frame.

The dual-stream network structure proposed in this paper can quickly extract the high-dimensional spatiotemporal features of the sequence and obtain the corresponding differential features based on the detected pedestrian area. Each convolution kernel is followed by a pooling operation. The pooling operation includes average pooling and maximum pooling. The adopted calculation formula is shown as

$$P_{i \rightarrow j} = \frac{(S_1 \oplus S_{i+1} \oplus \dots \oplus S_j)}{(j - i + 1)}, \quad (1)$$

where $P_{i \rightarrow j}$ represents the average pooling feature between key frames i to j . After the key frames are processed by convolution pooling and fully connected operation, the final output result of the deep difference network is a d -dimensional feature vector, and finally, the high-dimensional spatiotemporal information of the entire sequence is obtained. Each key frame is formed into a $1 \times 1 \times 1024$ -dimensional vector after the global average pooling operation, and then, the final spatiotemporal feature is extracted through the last convolutional layer.

3.2. Decision-Level Fusion Mechanism with Improved Loss Function. The spatiotemporal dual-channel branch of the dual-stream network separately extracts features from the same sequence of different modal images to obtain spatial position information and temporal motion information. These two types of features enhance the characterization

ability of pedestrian actions under the action of the fusion module, but the original dual-stream network only uses feature cascade for fusion. In addition, due to the complexity of the human action recognition problem in video sequential, its performance is often susceptible to interference from environmental noise and ultimately make wrong decisions and affect the output of the entire model. In order to improve the accuracy of the recognition model, this paper proposes a decision-level fusion mechanism. The fusion mechanism draws on the memory characteristics of the LSTM network, by modeling the previous output data and using a coupling mechanism to associate information in different dimensions. It has characteristic invariance in high-dimensional space.

A strong classifier based on improved Softmax logistic regression is designed in paper. The fused features are used to classify the pedestrian action and can get the highest classification probability, which can more effectively improve the accuracy of action recognition. In [15], assuming that the currently given sample sequences $x^{(i)}$ have k categories, the output is $y^{(i)} \in \{1, 2, \dots, k\}$, where its training set is $\{(x^{(i)}, y^{(i)})\}$, $i \in \{1, 2, \dots, k\}$. For a given sample feature x , the estimated conditional probability of the category j is $p(y = j|x)$, and the probability equation can be expressed as the following equation:

$$p(y^{(i)} = j | x^{(i)}; \theta) = \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}}. \quad (2)$$

Therefore, the classification probability of each category in Softmax logistic regression is written as follows:

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix}. \quad (3)$$

Since the probability of each category satisfies the exponential family distribution [18], if the recognition probability $h_{\theta}(x^{(i)})$ obtained by equation (3) is expanded in series, we can obtain

$$h_{\theta}(x^{(i)}) = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}. \quad (4)$$

The model parameter θ is a matrix with k rows, where each row represents the parameters of the corresponding category, so the model parameter matrix θ can be written as $\theta = [\theta_1^T, \theta_2^T, \dots, \theta_k^T]$. $1/\sum_{j=1}^k e^{\theta_j^T x^{(i)}}$ in equation (4) is the normalization operation on the probability distribution, so as to quantify the output probability. By deriving the log-likelihood of the overall sample results, the loss function is rewritten as follows:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^k I(y^{(i)} = j) \log \frac{e_j^T x^{(i)}}{\sum_{l=1}^k e_l^T x^{(i)}}, \quad (5)$$

where $I(y^{(i)} = j)$ is an indicative function. The value is 1 when there is a positive example, otherwise it is 0. In order to minimize the model parameter matrix θ , the obtained probability value by substituting equation (4) into equation (5) is expressed as

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^k y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log(1 - \log h_{\theta}(x^{(i)})). \quad (6)$$

In order to minimize the loss function shown in equation (6), the gradient descent method is generally used for optimization, and the partial derivative of the loss function is calculated as follows:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m (x^{(i)} I(y^{(i)} = j) - p(y^{(i)} = j | y^{(i)}; \theta)). \quad (7)$$

The first l partial derivative $\partial J(\theta)/\partial \theta_{jl}$ of the probability vector $\nabla_{\theta_j} J(\theta)$ indicates that the loss function takes the partial derivative of l parameters of category j . The gradient descent iteration of equation (8) is updated to determine the minimized loss function. The iterative operation includes the following equation:

$$\theta_j = \theta_j - \alpha \nabla_{\theta_j} J(\theta). \quad (8)$$

However, the update strategy of equation (8) used in Softmax logistic regression will affect the update effect of the parameters. Therefore, the multiobjective classification network, proposed in [16], is used to optimize the model, and the probability of equation (2) can be rewritten as $e^{(\theta_j - \varphi)^T x^{(i)}} / \sum_{j=1}^k e^{(\theta_j - \varphi)^T x^{(i)}}$, and the equation is expanded, and we get $p(y^{(i)} = j | x^{(i)}; \theta) = e^{\theta_j^T x^{(i)}} / \sum_{j=1}^k e^{\theta_j^T x^{(i)}}$. That is to say, when all the hyperparameters θ are subtracted from $e^{\varphi^T x^{(i)}}$, the probability value of the loss function does not change, which shows that when Softmax classifies different

samples, the result is not affected by the initial value, but this may cause the optimal solution to be nonunique. In order to solve this problem, this paper introduces a regular weight attenuation term λ in the loss function, constrains its optimal solution, and speeds up the convergence process. Therefore, the improved loss function in this paper can be rewritten as the following equation:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^k I(y^{(i)} = j) \log \frac{e_j^T x^{(i)}}{\sum_{j=1}^k e_j^T x^{(i)}} + \frac{\lambda}{2} \sum_{j=1}^k \sum_{j=0}^n \theta_{ij}^2. \quad (9)$$

When λ is greater than 0, the partial derivative of equation (9) is denoted as follows:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m [x^{(i)} I(y^{(i)} = j) - p(y^{(i)} = j | y^{(i)}; \theta)] + \lambda \theta_j. \quad (10)$$

In order to solve the above improved Softmax logistic regression equation and obtain the classification probability w_i of pedestrian actions, this paper constructs a decision-level fusion mechanism, which makes decisions on category probabilities w_i^f and w_i^d under different samples. For different action categories, the recognition probability p_k of the input image can be obtained by using the principle of multiplication.

$$p_k = \frac{P_{k,w^d} \times P_{k,w^f}}{\sum_{k=1}^k P_{k,w^d} \times P_{k,w^f}}. \quad (11)$$

Obtain the recognition probability of multiple images through equation (11), and find the maximum value from it as the final recognition probability u of the current image sequence:

$$u = \arg \max_{k,i} p_{k,i} \quad k = 1, 2, \dots, K, \quad (12)$$

where i is the number of videos included in each type of action and k is the total number of action types.

This paper proposes a cost function based on the decision-level feature fusion mechanism, which can retain the spatial and temporal information of images between different network on a larger scale, and adopts the principle of majority voting to increase the recognition probability of action categories under different key frame sequences, thereby improving the performance of human motion recognition.

4. Experimental Results and Analysis

4.1. Action Dataset. The scene image detected by the low-cost surveillance camera has no obvious texture details, and it is difficult to obtain fine behaviors such as playing ball and smoking through the image. For pedestrian detection and motion recognition tasks, most of the existing models use three datasets: OTCBVS, KAIST, and FLIR for pedestrian detection, but it is difficult to analyze pedestrian behavior. The main blame is that these images are not continuous sequences, and their movement time span is large, which

makes it difficult to conduct correlation analysis. The InfAR dataset [18] is currently a benchmark dataset publicly available in the field of behavior recognition, including 12 daily behaviors such as walking, fighting, clapping, shaking hands, jogging, and hugging. Each behavior type has 50 video sequences. It is completed by a single person or multiple people interactively, but the amount of data is limited. Most algorithms perform transfer learning on visible datasets to improve the recognition effect of behavior in sequences.

The model proposed in this paper mainly analyzes the behavior characteristics of a single pedestrian in the surveillance area. Therefore, the project team collected a large number of pedestrian motion videos to help improve the performance of the model. In order to facilitate performance comparison, this paper also established a self-built dataset, and these pedestrians and their actions have been annotated in the image, including standing, squatting, lying, running, and other action categories. The number of all categories is relatively balanced with a total of 3115 video clips. Table 1 shows the number of sequences in different categories. The first 12 categories are single-person actions, and the last four categories are multiperson interactive actions.

4.2. Parameter Setting. The resolution of all the images in this paper is 6400×512 , and the performance of the proposed model is verified by 5-fold cross validation. The networks selected in this paper are all based on TensorFlow framework. The random gradient descent method is used to learn the network parameters. The batch size is 128, and the momentum value and weight attenuation are set to 0.9 and 0.0005, respectively. The initial value of learning rate is 0.01. The learning rate remains unchanged at 0.001 during the first 50 rounds in process of training, and then, the learning rate will be reduced by 10% every 10 rounds so as to prevent overfitting. The value range of λ is $\{1e-3, 1e-2, 1e-1, 1, 1e1, 1e2, 1e3\}$.

In order to objectively analyze the effectiveness, the precision rate (PR), miss rate (MR), and recall rate (RR) are selected to quantify the detection performance. All the indicators can be calculated by true positive (TP), false positive (FP), false negative (FN), and true negative (TN). In addition, the confusion matrix is also used to analyze the effectiveness.

4.3. Ablative Analysis. In this paper, an improved pedestrian action recognition model based on deep differential dual-stream network is proposed. In this model, deep differential network is used instead of temporal network to obtain the difference map of feature propagation based on the key frame, which can reduce the computational complexity and enhance the representation ability of feature propagation map of human motion. At the same time, the strong classifier based on improved Softmax logistic regression is used to make pedestrian behavior category decision to improve the ability of category recognition. In order to analyze the effects of different improvement measures, an ablation analysis will be performed in this section. Table 2 shows the

TABLE 1: Categories and number of datasets.

No.	Categories	Total
1	Walking	152
2	Standing	203
3	Climbing	186
4	Jogging	265
5	Jumping	174
5	Punching	128
7	Lying	295
8	Waving 1	168
9	Waving 2	177
10	Crouching	312
11	Sitting	268
12	Handclapping	208
13	Pushing	158
14	Fighting	119
15	Handshaking	134
16	Hugging	168

TABLE 2: Performance analysis of different modules.

DDN	IS	DF	Pr (%)	FPS
			77.12	13.9
✓			77.83	18.1
	✓		79.91	13.8
		✓	79.78	12.7
✓	✓		81.79	17.8
✓		✓	82.09	18.5
	✓	✓	81.83	11.6
✓	✓	✓	83.01	17.7

recognition effect of deep differential network (DDN), improved Softmax (IS), and decision fusion (DF) on pedestrian motion sequence. The box \checkmark indicates the substituted modules in the benchmark network.

The first line in Table 2 does not replace any modules, which is the original dual-stream network. The recognition accuracy and frame rate are 78.12% and 13.9%, respectively. In ablation analysis, if different modules are replaced, their performance will be changed accordingly. Especially, the feature propagation differential map can replace the complex optical-flow calculation, which can greatly improve the processing efficiency. The frame rate is increased by 17.1%, and the accuracy is also improved. The improvement of IS and DF can also improve the performance and increase the recognition accuracy by 3% and 1.1%, respectively. If two modules are replaced in the original dual-stream network at the same time, it can be seen that the performance of any module is better than that of only one module. It is worth noting that the three improved modules designed in this paper are mainly to optimize the efficiency and accuracy. As long as they are replaced by the deep differential network, the final recognition efficiency will be greatly improved. The main reason is that the feature propagation differential map has the advantages of fast generation speed and small computational complexity. Finally, the three modules improve the dual-stream network from different angles and achieve 82.01% recognition accuracy and 17.7 processing frame rate for the video sequences.

4.4. Qualitative and Quantitative Analysis. In order to analyze the performance of the proposed pedestrian behavior recognition algorithm, we select the common behavior recognition algorithms for performance comparison, which are IDT (improved density trajectories) [19], C3D (continuous 3D) [20], SCNN-3G [21] (spatiotemporal continuous neural network based on 3D gradients), L-LSTM [22] (lattice long short-term memory), Ts-3D [23] (two-stream expanded 3D convolutional), and OFGF [24] (optical flow guided feature), where IDT is a very classic traditional algorithm in the field of behavior recognition. By introducing the background optical-flow elimination method and extracting features along the trajectory, the obtained features are more suitable for the description of human motion. C3D is to construct three-dimensional multichannel convolution features for continuous frames, extract multidimensional features through prior knowledge, and enhance the training speed and feature representation ability of back propagation. L-LSTM is a behavior recognition model based on rasterized long-term and short-term memory, which acts on video sequence by convolution and assumes that the motion in the video is stationary in different spatial positions. Ts-3D is a behavior recognition algorithm based on improved dual-stream network, which is extended from 2DCNN Inception-V1, and can use pretraining parameters to enhance the training efficiency. OFGF is a fast and robust motion representation method for video motion recognition. It can obtain the human motion trend by calculating the spatio-temporal gradient. All the comparison algorithms are tested using the author's source code. Because part of the original code is mainly for 3D natural image analysis, the research object of this paper is the 2D gray image. For the consistency of the algorithm model, all input images are expanded into three channel images. In addition, all experiments in this paper use the same test set and training set for comparison.

Because the digital video output of the camera reaches 100 frames, the content of adjacent frames changes very slowly. In order to make the input sequence fully extract temporal information, this paper uses multiscale frame extraction strategy to obtain the input dataset, which ensures that the input sequence can obtain more abundant temporal information on the premise of fixed dimension. Therefore, the data between some key frames is very redundant, and only a small amount of information is needed to represent the human movement trend. That is to say, the trend information independent of the duration can be obtained by using the differential key frame to ensure that the obtained feature information is evenly distributed along the time dimension. It can be seen that the improved strategy in this paper has the similar performance of optical-flow map and can fully represent the temporal action information of human behavior, but the computational complexity is smaller.

Figure 4 shows the differential key frame and the corresponding optical-flow map. It is obvious that most of the background noise in the sequence has been deleted, and the human action has been successfully retained. In addition, the differential information obtained in this paper is similar to the optical-flow information of the original image. This

operation can not only reduce the computational complexity but also make the model more robust.

Figures 5 and 6 show the loss value and recognition accuracy in the training process for the proposed model. In the training process, the learning rate of the loss function is updated dynamically with the change of the number of training rounds to prevent the training process from overfitting. It can be seen from the results in Figure 5 that the loss convergence is faster and more stable after using the decision fusion mechanism, while the training loss without fusion is more jittery. In Figure 6, the training accuracy of the fused dual-stream network can rise rapidly and approach 99%, which indicates that the fusion mechanism can effectively fuse the spatial-temporal information, and improve the representation ability of human action through complementary feature information.

It is very difficult to subdivide all the categories because of the diversity of human behavior. This paper mainly verifies the performance of the proposed algorithm. Therefore, 16 kinds of behaviors, such as standing, walking, running, and jumping, were selected for recognition. Yolo-v3 is used for the human detection model. The differential map obtained in this paper is refined on the basis of pedestrian detection results, and the range of processing is reduced, which helps to improve the representation ability of human movement trend. Table 3 shows the quantitative results of all comparison algorithms under the same test set. It can be seen that although the results of IDT in some datasets are not as good as the behavior recognition algorithm based on deep learning, and the performance can be improved by integrating the results of IDT in general, especially in the standing sequence. C3d and Ts-3D (two-stream 3D) are the two mainstream methods for behavior recognition, and their recognition accuracy reaches 75.2%. However, these two methods rely heavily on the variation of adjacent time sequences. Once there are fewer frame sequences between the key frames, their performance will be greatly reduced. For example, large amplitude motion leads to large variation of adjacent frames, and the final recognition accuracy is insufficient. For example, the result of sequence 2 is only 57%. L-LSTM often relies on the last layer feature of convolution network as input, which cannot capture low-level motion features, and it is difficult to train for traversing the whole video. In order to improve the long-time behavior recognition, dense sampling is a common method to improve the performance, but it requires huge computational overhead. The accuracy rate, missed detection rate, and recall rate of OFGF are 73.8%, 19.2%, and 78.4%, respectively. Although this is the best algorithm of the contrast algorithm, it can be embedded in any existing deep network framework with only a small time cost, and its processing frame rate reaches 69.7. The model proposed in this paper inputs it into the deep differential network to extract the time dimension features, which can ensure the accuracy and greatly reduce the operation time, and finally get 78% recognition accuracy. It can be seen that the recognition accuracy of this model is 6.7% higher than that of L-LSTM, and 1.8% higher than that of C3D with 68 layers,

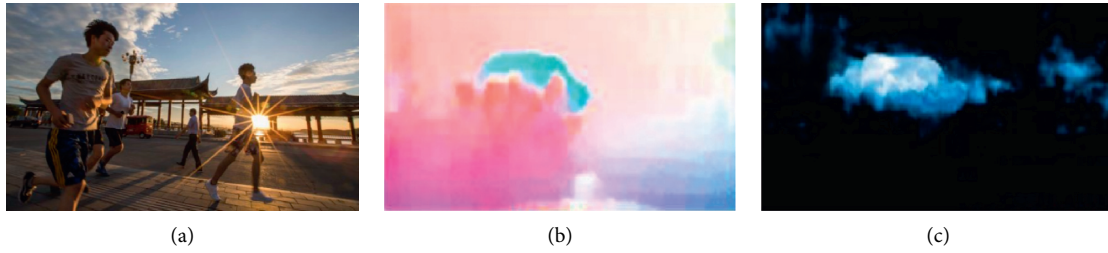


FIGURE 4: Differential key frame and corresponding optical flow. (a) Key frame; (b) optical flow; (c) differential key frame.

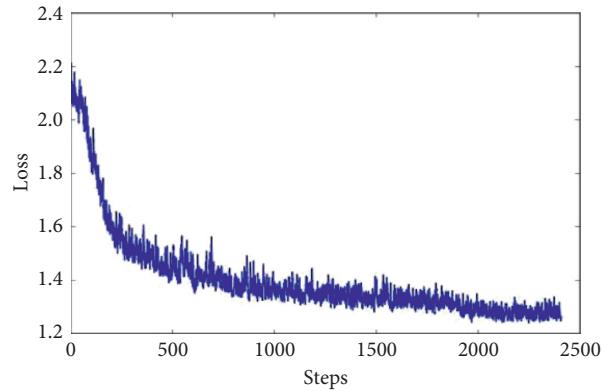


FIGURE 5: Change trend of loss value during training.

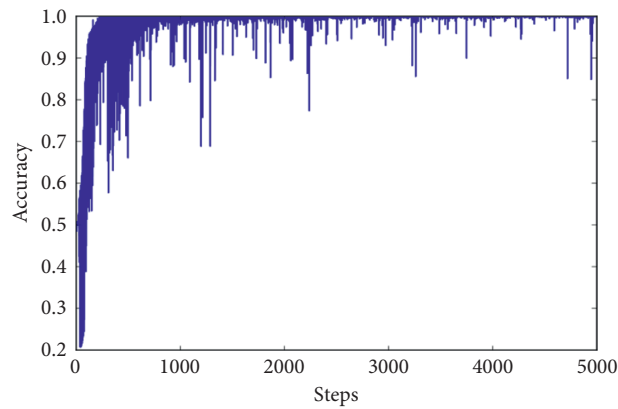


FIGURE 6: Change trend of precision during training.

which fully shows that the proposed model can recognize human actions more effectively. Figure 7 shows the histogram of the recognition results of comparison algorithms on the mixed data set, so as to intuitively analyze the performance of the proposed model, where the ordinate is the percentage of different index results. From the histogram trend, it can be seen that the proposed model has better results in the quantitative evaluation index precision rate (PR), miss rate (MR), and recall rate (RR).

To analyze the recognition performance of the proposed method for different types of behaviors, we also analyze the confusion matrix before and after the fusion strategy. According to the experimental results, walking and standing can achieve a high recognition rate whether or not using this method. When using the decision fusion method, the recognition rate of hugging, shaking hands, fighting, and other behaviors is slightly low, and there are many wrong points. Using our proposed network can improve the wrong points between these behaviors

TABLE 3: Performance analysis of different comparison models

Categories	IDT			C3D			SCNN-3G			L-LSTM			Ts-3D			OFGF			Our		
	Pr	Mr	Rr	Pr	Mr	Rr	Pr	Mr	Rr	Pr	Mr	Rr	Pr	Mr	Rr	Pr	Mr	Rr	Pr	Mr	Rr
Walk	65	28	71	67	22	73	69	24	73	75	20	78	77	28	75	80	17	81	79	11	81
Stand	73	21	76	77	20	78	77	20	75	83	20	88	85	21	76	85	17	86	86	21	87
Climb	51	37	62	54	32	64	62	35	67	67	26	68	72	37	62	77	25	82	79	17	82
Jog	67	29	71	69	24	76	71	24	71	68	29	77	72	29	71	77	20	79	87	9	91
Jump	61	33	66	62	32	69	68	35	68	61	33	75	73	33	66	73	23	78	72	17	81
Punch	42	51	45	42	41	44	47	52	49	52	41	59	61	51	65	62	31	65	68	23	70
Lying	57	37	61	58	32	67	60	34	66	57	37	68	71	31	68	67	23	70	68	17	71
Wave 1	66	32	66	69	30	69	69	31	69	66	32	77	73	24	76	76	12	81	83	12	86
Wave 2	69	29	70	71	31	72	72	24	77	69	29	88	79	29	80	82	18	87	89	9	89
Crouch	42	30	42	44	35	46	45	24	47	42	30	59	54	21	51	61	23	62	69	27	72
Sitting	71	25	79	74	29	81	73	29	80	72	25	82	79	20	82	81	16	89	83	15	88
Handclap	38	34	39	39	35	43	39	31	34	38	34	51	46	24	59	68	23	69	73	24	77
Push	42	47	45	45	48	47	43	43	48	42	47	58	67	31	65	72	24	75	72	17	80
Fight	54	36	58	59	31	59	57	32	59	54	36	68	68	30	68	64	16	78	81	14	81
Handshake	63	30	68	66	32	71	67	27	71	63	30	77	72	21	78	76	20	88	77	23	82
Hug	68	27	70	67	28	73	62	29	75	77	29	75	75	27	79	79	26	80	82	15	86
Mixed	58	32	61	60	31	64	61	30	64	61	31	71	70	28	70	73	19	78	78	16	81

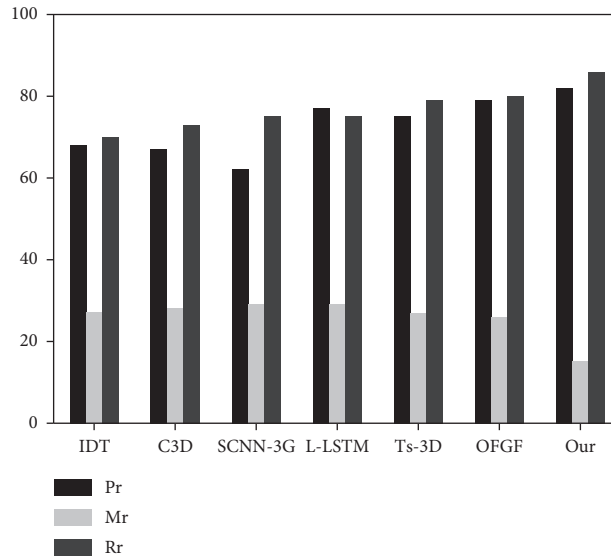


FIGURE 7: Qualitative analysis of different models under mixed datasets.

and improve the correct recognition rate. The simulation results show that the proposed improved network achieves 87% recognition accuracy on the self-built infrared dataset, and the computational efficiency is improved by 15.1%.

5. Conclusion

Aiming at the problem of low accuracy of pedestrian behavior recognition in image sequences, this paper proposes a pedestrian action recognition model based on improved spatial-temporal dual-stream network. The model uses differential key frames instead of optical flow sequence for temporal feature extraction, which ensures the accuracy and greatly reduces the computational complexity. At the same time, this paper also uses the cost function based on the decision-level feature fusion

mechanism to train the model, which can retain the spatio-temporal characteristics of images between different network frames to a greater extent, and reflect the action category of pedestrians more realistically. Simulation experiments also verify the effectiveness of the model from different angles. The next step is to refine the behavior categories, establish a larger and richer training sample set, improve the recognition accuracy and generalization ability of the model, and transplant the model on the basis of AI embedded platform to realize behavior recognition in complex monitoring environment.

Data Availability

The labeled dataset 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

This paper was supported by the construct program of applied characteristic discipline in Hunan University of Science and Engineering.

References

- [1] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. F. Fei, "Large-scale video classification with convolutional neural networks," in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1725–1732, Columbus, OH, USA, June 2014.
- [2] D. Tran, L. D. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3d convolutional networks," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 4489–4497, Santiago, Chile, December 2015.
- [3] B. Zhang, L. Wang, Z. Wang, Y. Qiao, and H. Wang, "Real-time action recognition with enhanced motion vector CNNs," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2718–2726, Las Vegas, NV, USA, June 2016.
- [4] J. C. Niebles, C. W. Chen, and L. F. Fei, "Modeling temporal structure of decomposable motion segments for activity classification," in *Proceedings of the European Conference on Computer Vision ECCV*, pp. 392–405, Heraklion, Crete, Greece, September 2010.
- [5] P. Tumas, A. Nowosielski, and A. Serackis, "Pedestrian detection in severe weather conditions," *IEEE Access*, vol. 8, Article ID 62775, 2020.
- [6] W. Li, M. Ding, and L. Zeng, "Pedestrian detection based on objectness and sparse coding in a single infrared image," *Infrared Technology*, vol. 38, no. 9, pp. 752–757, 2016.
- [7] B. Fernando, E. Gavves, O. M. Jose, A. Ghodrati, and T. Tuytelaars, "Modeling video evolution for action recognition," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5378–5387, Boston, MA, USA, June 2015.
- [8] G. Varol, I. Laptev, and C. Schmid, "Long-term temporal convolutions for action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, Article ID 04494, 2018.
- [9] J. Donahue, L. H. Anne, and S. Guadarrama, "Long-term recurrent convolutional networks for visual recognition and description," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2625–2634, Boston, MA, USA, June 2015.
- [10] K. Soomro, A. R. Zamir, and M. Shah, "UCF101: a dataset of 101 human actions classes from videos in the wild," *CoRR abs*, vol. 1212, 2012.
- [11] H. Kuehne, H. Jhuang, E. Garrote, T. A. Poggio, and T. Serre, "HMDB: a large video database for human motion recognition," in *Proceedings of the 13th International Conference on Computer Vision*, pp. 2556–2563, Barcelona, Spain, November 2011.
- [12] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the The 32nd International Conference on Machine Learning (ICML)*, pp. 448–456, Lille, France, July 2015.
- [13] L. Wang, Y. Qiao, and X. Tang, "Video action detection with relational dynamic-p," in *Proceedings of the 13th European Conference on Computer Vision ECCV*, pp. 565–580, Zurich, Switzerland, September 2014.
- [14] C. Gan, T. Yao, K. Yang, Y. Yang, and T. Mei, "You lead, we exceed: labor-free video concept learning by jointly exploiting web videos and images," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 923–932, Las Vegas, NV, USA, June 2016.
- [15] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," *Advances in Neural Information Processing Systems*, vol. 150, pp. 109–125, 2014.
- [16] R. Peng, L. Wang, L. Xin, and P. Liu, "Deep convolution neural network based on improved softmax classifier and its application in face recognition," *Journal of Shanghai University (Social Sciences Edition)*, vol. 24, no. 3, pp. 352–366, 2018.
- [17] H. Yasin, M. Hussain, and A. Weber, "Keys for action: an efficient keyframe-based approach for 3D action recognition using a deep neural network," *Sensors*, vol. 20, no. 8, p. 2226, 2020.
- [18] A. G. Hauptmann, Y. Du, J. Liu, J. Lv, L. Yang, D. Meng et al., "InfAR dataset: infrared action recognition at different times," *Neurcomputing*, vol. 212, pp. 36–47, 2016.
- [19] H. Wang and C. Schmid, "Action recognition with improved trajectories," in *Proceedings of the 2013 IEEE International Conference on Computer Vision*, pp. 3551–3558, IEEE, Sydney, Australia, December 2013.
- [20] D. Tran, L. Bourdev, and R. Fergus, "Learning spatiotemporal features with 3D convolutional networks," in *Proceedings of the 2015 IEEE International Conference on Computer Vision*, pp. 4489–4497, IEEE, Boston MA, USA, June 2015.
- [21] T. Yang, Z. Chen, and W. Yue, "Spatio-temporal two-stream human action recognition model based on video deep-learning," *Journal of Computer Applications*, vol. 38, no. 3, pp. 895–899, 2018.
- [22] L. Su, K. Jia, K. Chen, D. Y. Yeung, B. E. Shi, and S. Silvio, "Lattice long short-term memory for human action recognition," in *Proceedings of the 2017 IEEE International Conference on Computer Vision*, pp. 2166–2175, IEEE, Venice, Italy, October 2017.
- [23] J. Carreira and A. G. S. R. Man, "Quo vadis, action recognition? A new model and the kinetics dataset," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4724–4733, IEEE, Honolulu, HI, USA, July 2017.
- [24] S. Sun, Z. Kuang, L. Sheng et al., "Optical flow guided feature: a fast and robust motion representation for video action recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20118–20132, IEEE, Salt Lake City, Utah, USA, June 2018.

Research Article

A Novel Brain Image Segmentation Method Using an Improved 3D U-Net Model

Zhuqing Yang 

Jiangsu Vocational College of Information Technology, Wuxi, Jiangsu 214153, China

Correspondence should be addressed to Zhuqing Yang; yangzq@jsit.edu.cn

Received 26 July 2021; Revised 10 August 2021; Accepted 13 August 2021; Published 19 August 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Zhuqing Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical image segmentation (IS) is a research field in image processing. Deep learning methods are used to automatically segment organs, tissues, or tumor regions in medical images, which can assist doctors in diagnosing diseases. Since most IS models based on convolutional neural network (CNN) are two-dimensional models, they are not suitable for three-dimensional medical imaging. On the contrary, the three-dimensional segmentation model has problems such as complex network structure and large amount of calculation. Therefore, this study introduces the self-excited compressed dilated convolution (SEDCD) module on the basis of the 3D U-Net network and proposes an improved 3D U-Net network model. In the SEDCD module, the calculation amount of the model can be reduced by $1 \times 1 \times 1$ convolution. Combining normal convolution and cavity convolution with an expansion rate of 2 can dig out the multiview features of the image. At the same time, the 3D squeeze-and-excitation (3D-SE) module can realize automatic learning of the importance of each layer. The experimental results on the BraTS2019 dataset show that the Dice coefficient and other indicators obtained by the model used in this paper indicate that the overall tumor can reach 0.87, the tumor core can reach 0.84, and the most difficult to segment enhanced tumor can reach 0.80. From the evaluation indicators, it can be analyzed that the improved 3D U-Net model used can greatly reduce the amount of data while achieving better segmentation results, and the model has better robustness. This model can meet the clinical needs of brain tumor segmentation methods.

1. Introduction

A basic task in medical IS [1, 2] is to extract specific organs and tumors from different types of medical images. Organ and tumor segmentation provides an important basis for cancer diagnosis, surgical planning, and pathological analysis. The segmentation of organs and tumors in the clinic faces two main challenges. The first challenge is the quality of medical images. Due to the diversity of clinical needs, doctors usually choose different imaging examination methods to produce different types of medical images, such as computed tomography (CT) [3], magnetic resonance imaging (MRI) [4], X-ray [5], and ultrasound [6]. Different types of imaging equipment have different principles and will be interfered by various factors during the image acquisition process. There are many reasons for interference, which can be summarized into the following three points.

One is that due to individual differences, the organs or tumors of patients of different body shapes and ages will show different shapes of anatomical structures on medical images. Second, in the image acquisition process, the patient's breathing, blood flow, heart circulation, and other factors will cause the anatomical structure to deform. The overlapping of soft tissues such as organs causes unclear boundaries. The third is the defects of the equipment itself and the interference from the outside. The above factors will reduce the image quality and increase the difficulty of medical IS. The second challenge is the characteristics of the disease itself. Some tumors appear in the early stage of medical images with small volume, small difference in texture from surrounding tissues, and low contrast, which affects the accuracy of segmentation. These small tumors are easily misdiagnosed as calcification or fatification or even ignored, leading to misdiagnosis or missed diagnosis of

cancer. It is an arduous and complicated task to screen out images with tumors and segment them from a large number of medical images. Therefore, there is an urgent clinical need to develop a method that can accurately and automatically segment tumor regions from a large number of medical images to assist doctors in cancer diagnosis.

Traditional medical IS methods include threshold method [7, 8], graph cut method [9, 10], and region growing method [11, 12]. The threshold method mainly calculates the threshold value suitable for the segmentation task according to the manually designed criteria and compares the gray value of each pixel in the image with the threshold value, thereby separating the target from the background. The core idea of the region growing method is to select a seed pixel in different regions and use the texture, gray, gradient, color, and other characteristics as the criterion to measure the similarity between all other pixels in the region and the seed pixel. Pixels that meet the similarity criterion are classified into one category to achieve IS. The core idea of the graph cut method is to take the pixels of the image as the vertices of the graph and establish the graph structure from the predefined vertex connection relationship. By designing a suitable cutting criterion to remove the edges that do not conform to the criterion, several unconnected subgraphs are obtained. These subgraphs constitute the final segmentation result. Traditional IS methods rely on manually set parameters, and setting these parameters requires a lot of medical expertise. With the emergence of bottlenecks in traditional medical segmentation methods, there are more and more medical IS studies based on deep learning methods. The segmentation model based on deep learning does not require manual setting of parameters. Under the guidance of supervision information, features can be automatically learned from the given training samples, which significantly improve the efficiency and accuracy of IS.

In recent years, medical IS methods based on deep learning have made good progress. All fully connected layers of neural networks such as AlexNet [13] and VGG Net [14] are replaced with convolutional layers, and fully convolutional networks (FCNs) [15] are proposed, and this network is applied to the field of IS for the first time. In order to extract multiscale features of an image, convolution or pooling operations are usually used to change the feature size of the image. Since the size of the IS result needs to be consistent with the original image size, an upsampling operation is used to restore the feature map size to the original image size. The part of the network used to extract features is usually called the encoder, and the part of the network that is restored from the features to the original image size is called the decoder. The encoder and decoder form an encoding-decoding structure. Based on the encoding-decoding structure, U-Net [16] introduces skip connections to integrate low-level semantic information with high-level semantic information, which further improves the segmentation performance of the network. The success of U-Net made other networks choose U-Net as the backbone of the model. Iglovikov and Shvets [17] regard VGG Net as the encoder of U-Net and improve the performance of U-Net by pretraining the weights of VGG Net

on ImageNet. Attention U-Net [18] introduces an attention mechanism into the decoder of U-Net [19], which effectively suppresses the influence of areas that are not related to the target in the medical image. Sun et al. [20] introduce the Attention-Up module to improve the symmetric structure of U-Net to an asymmetric structure.

The above methods mainly use 2D models to segment clinical medical images. However, clinical medical images are usually stacked by multiple slices, and adjacent slices of the image are sometimes related to each other. Only the 2D model cannot learn the features between image slices. With the continuous deepening of research, many 3D segmentation models for 3D medical images have begun to emerge. V-Net [21] and 3D U-Net [22] replace two-dimensional convolution with three-dimensional convolution and realize the transformation from the 2D model to the 3D model. Compared with the 2D model, the 3D model encodes the image in three directions and extracts the three-dimensional features of the image. 3D models need to consume a lot of computing and storage resources when calculating, so the image needs to be cropped into image blocks of a certain size and input into the network for calculation. In order to take advantage of the respective advantages of the 2D and 3D medical IS models, H-Dense U-Net [23] combines the 2D segmentation model with the 3D segmentation model. The 2D model is used to extract the intraslice features of the three-dimensional medical image, and the 3D model is used to extract the interslice features, and the segmentation accuracy is improved by fusing the intraslice features and the interslice features.

Although the CNN-based IS model performs well in a variety of segmentation tasks, its performance still has room for improvement. Since CNN-based IS models are mostly two-dimensional models, they are not suitable for three-dimensional medical imaging. Therefore, this article uses 3D U-Net model to segment medical images. The work done in this study is summarized as follows:

- (1) The advantages and disadvantages of various network models in medical IS are compared and analyzed, and the 3D U-Net network is used for brain IS tasks.
- (2) Due to the complex network structure and large amount of calculation in the 3D segmentation model, this paper introduces the SECDC module optimization model. The introduction of the SECDC module can reduce the calculation parameters and effectively reduce the amount of calculation.
- (3) The abovementioned improved 3D U-Net network is applied to the brain image dataset. The IS evaluation index is used to verify the experimental results, and the results show the effectiveness of the model used. Based on the better segmentation performance of the used model, it has certain clinical significance for the diagnosis of diseases.

2. D U-Net Model

2.1. Network Structure. At present, there are two strategies for 3D medical IS tasks. One is to send the 2D slices of the volume into the 2D network for training. The training speed

of this strategy is fast, but the spatial location information is insufficient. The second is to send the volume directly into the 3D network for training. This method has a large amount of parameters, a long training time, and high requirements on hardware conditions, but its segmentation effect is better than that of a 2D network. The use of CNN network can make 2D biomedical IS reach the accuracy close to human manual segmentation. It is because of such successful applications that people begin to use CNN network to segment 3D data. In 2016, Çiçek et al. [24] proposed a 3D U-Net network structure for learning 3D segmentation from sparsely annotated stereo data. The 3D U-Net model structure is shown in Figure 1.

2.2. Network Advantages and Disadvantages. The biggest feature of the 3D U-Net network is the U-shaped codec structure and jump connection so that the shallow features can be well integrated with the high-level abstract features. These features are very effective for medical images with continuous structure, fuzzy boundaries, and simple semantics. Without considering the calculation and memory performance, the 3D U-Net network can combine the information between image layers to ensure the continuity of changes between the image masks of the interlayer. This feature is easier to obtain better results than using 2D slices for training.

The network structure of 3D U-Net [25] realizes the segmentation in two scenarios. One is for semiautomatic segmentation, and the other is for fully automatic segmentation. In semiautomatic segmentation, users mark some slices in the volume to be segmented. The network then learns from these sparse annotations and provides a dense 3D segmentation result. In fully automated segmentation, it is assumed that there is a representative training set with sparse annotations. Trained on this dataset, the network can intensively segment new volumetric images.

The main disadvantage of the 3D U-Net network lies in the memory usage, which makes it impossible to use the entire 3D patch as input. Therefore, it needs to be tailored. Generally, the entire volume is cut into a series of 3D patches of the same size as input. Using patch for training will limit the size of the maximum receptive field that the network can reach, resulting in the loss of certain global information. If the target to be segmented is much larger than the patch block, it is difficult for the network to learn the overall structure of the target.

3. Image Segmentation Based on Improved 3D U-Net Model

3.1. Model Segmentation Process. Generally speaking, the IS process based on the network model is shown in Figure 2. First, preprocess the image to be segmented. Preprocessing mainly includes cutting a series of standardized processing. Second, the training data is input into the model structure, and the IS model is obtained through training. In this study, we use an improved 3D U-Net model for training. Third, when the network training is over, the performance of the

trained network needs to be tested. By inputting the test data, the evaluation index is calculated according to the test result. Multidimensional evaluation of trained network performance based on different evaluation indicators is carried out.

3.2. Improved 3D U-Net Model. This article introduces the SECDC module to optimize the convolution operation in 3D U-Net. This module first uses $1 \times 1 \times 1$ convolution to reduce the dimensionality of the network layer. Then, feature mapping is performed on the input layer with normal convolution and dilated convolution with an expansion rate of 2. Then, learn the importance of different layers through the 3D-SE module, and finally, use $1 \times 1 \times 1$ convolution to upgrade the dimension. The structure of the module is shown in Figure 3.

3.2.1. 3D-SE Module. The Squeeze-and-Excitation (SE) module [26] is mainly used to measure the relationship between channels so that the model can automatically learn the importance of different channel features. Two key operations, Squeeze and Excitation, are included in the SE module. The Squeeze operation aggregates the feature maps obtained by convolution to obtain the feature map with dimension $W \times H$ as the feature descriptor, so as to obtain the information of the global receptive field. The Excitation operation is a self-screening mechanism that uses a sample specialized activation function to evaluate the weights of all channels. The module structure is shown in Figure 4.

The mapping rules of the SE module can be expressed as $F_{tr}: X \rightarrow U, X \in R^{H' \times W' \times C'}, U \in R^{H \times W \times C}$. In the convolutional map, let the convolution kernel be $V = [v_1, v_2, \dots, v_c]$, where v_c represents the c th convolution kernel, and the output $U = [u_1, u_1, \dots, u_c]$ can be expressed as

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s, \quad (1)$$

where $*$ represents convolution and v_c^s represents a 3D convolution kernel. It inputs spatial features on a channel to learn the relationship between feature spaces. However, the convolution results of each channel are summed so that the feature relationship of each channel is merged with the spatial relationship of the convolution kernel. This fusion is not good for the training of the model. The SE module avoids this unprofitable fusion and aims to allow the model to directly obtain the characteristic relationship of each channel.

Compared with the original SE module, the improved 3D SE module can be applied to three-dimensional convolution. It mainly focuses on the importance of different channel features in three-dimensional space features. Figure 5 gives the structure of the 3D SE module.

The output formula of the 3D SE module is as follows:

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s, \quad (2)$$

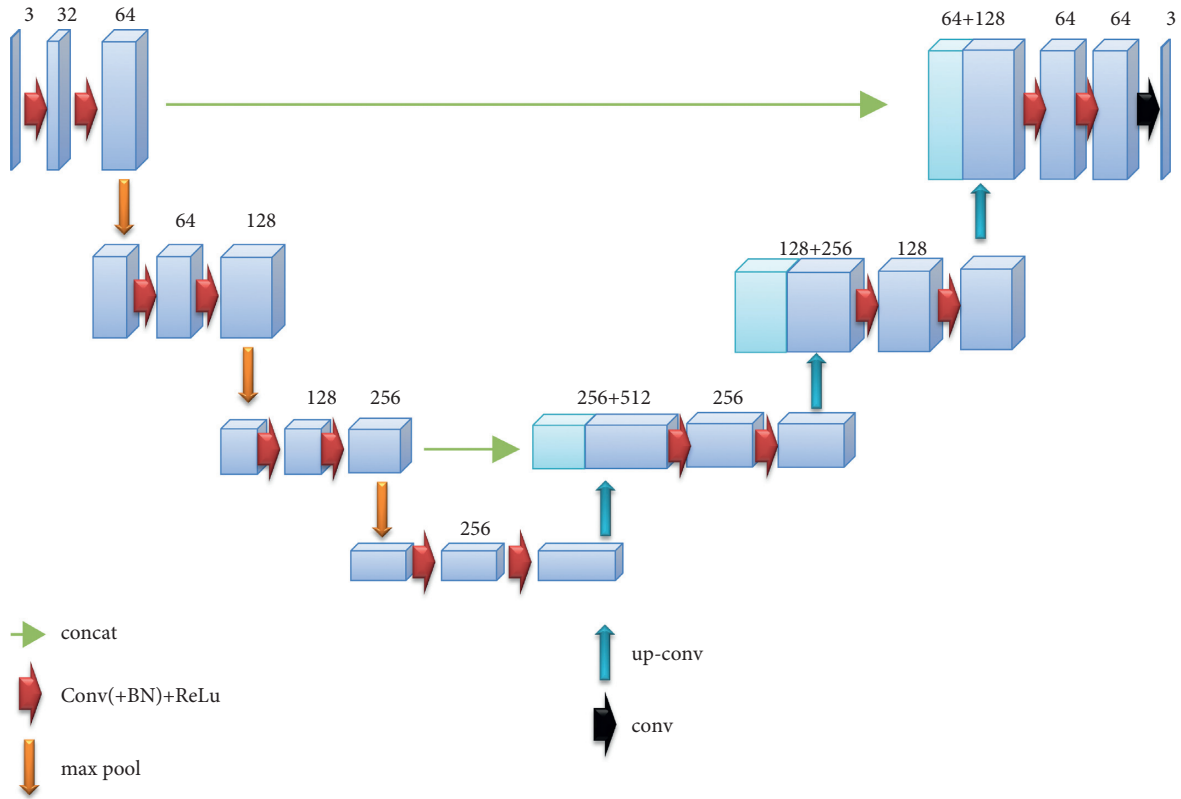


FIGURE 1: 3D U-Net model structure.

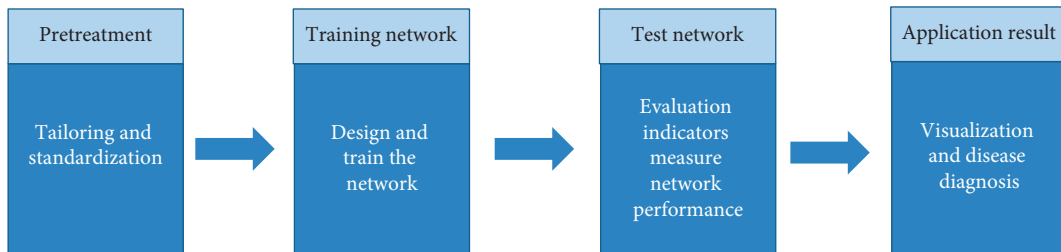


FIGURE 2: Image segmentation process based on the network model.

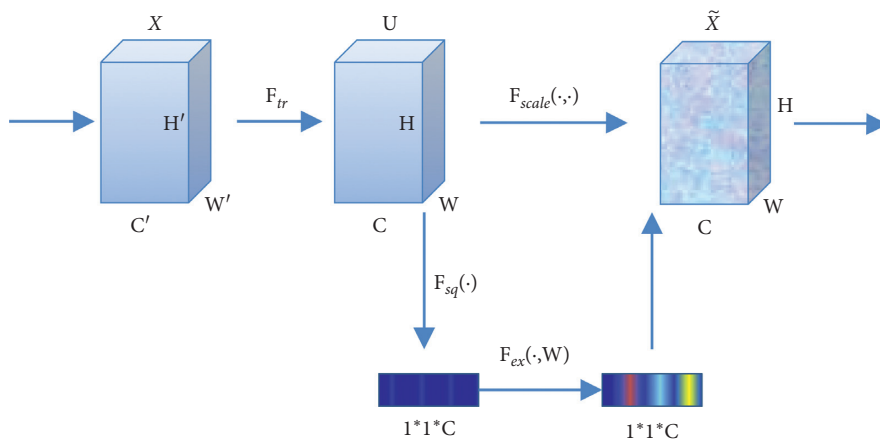


FIGURE 3: SECD module structure.

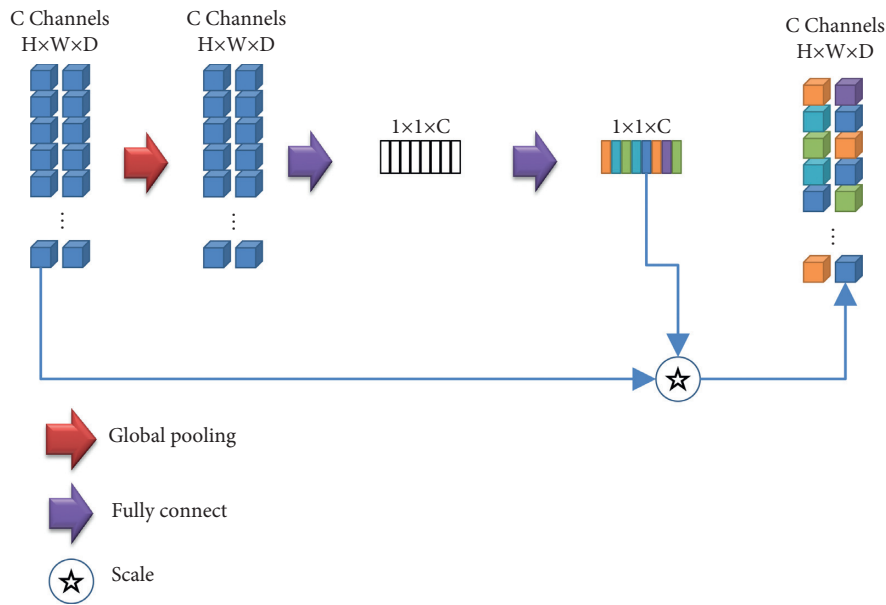


FIGURE 4: SE module structure.

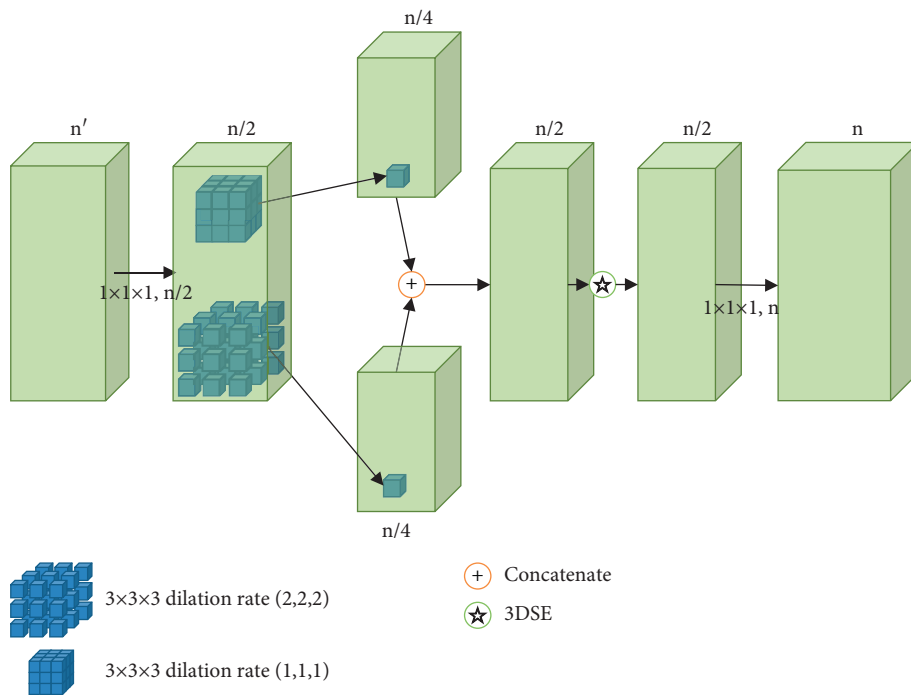


FIGURE 5: 3D SE module.

where the difference between SE and 3D SE modules is that γ_c^s represents a 4D convolution kernel, which can be directly combined with the 3D convolution operation in 3D U-Net.

3.2.2. $1 \times 1 \times 1$ Convolution. Add $1 \times 1 \times 1$ convolution before and after convolution, as shown in Figure 3. $1 \times 1 \times 1$ convolution has two main functions. One is to realize

cross-channel interaction and information integration. The $1 \times 1 \times 1$ convolution keeps the width, height, and depth of the data unchanged and linearly combines multiple features to achieve cross-channel information integration and improve the expressive ability of the network. The second is to reduce and increase the number of channels of the convolution kernel. Since the $3 \times 3 \times 3$ convolution is very time-consuming to perform convolution operations on the convolutional layer of hundreds of filters, the $1 \times 1 \times 1$

convolution is used to reduce and increase the dimensionality before and after the $3 \times 3 \times 3$ convolutional layer. Dimensional operation reduces the number of parameters and shortens time.

3.2.3. Dilated Convolution. Dilated convolution [27, 28] injects holes in the standard convolution map. Compared with normal convolution, the dilated has one more hyperparameter called the dilation rate, which represents the number of intervals between the convolution kernels. The dilatation rate of the convolution kernel for normal convolution operation is 1. The advantage of dilated convolution is that, without pooling loss information, the receptive field is enlarged so that each convolution output contains a larger range of information. The SECDC module structure is shown in Figure 3.

4. Experiment

4.1. Segmentation Tasks and Experimental Datasets. The segmentation task is to accurately segment the three tumor regions of brain tumor patients, namely, whole tumor (WT), tumor core (TC), enhanced tumor (ET), and background region. Among them, WT refers to the area containing all tumor types, TC refers to the area containing all gangrene types, and ET refers to the area containing all enhanced tumor types. The dataset used in the experiment is BraTS2019 [29]. The experimental dataset selected MRI images of 248 HGG patients and 68 LGG patients. The MRI image of each patient includes four registered modal images and real segmentation label images. Each modal and real segmentation label has 155 scanned images of different layers with a size of 240×240 .

Before network training, preprocess the MRI image. Cut and standardize the image first. The size of the processed image is 160×160 , and methods such as flipping, rotation, and elastic deformation continue to be used for data enhancement. The test data enhancement here refers to multiple transformations of an image, including folding, rotating, cropping, scaling, and adding random noise. Input it into the model and calculate multiple versions of data and finally get the average output as the final output of the image. Use 5-fold cross validation in the experiment to avoid data bias. Images of 180 HGG patients and 50 LGG patients were used as training samples. Images of 68 HGG patients and 18 LGG patients were used as the test set. The network inputs the data of each patient, including preprocessed MRI images of FLAIR, T1, T1ce, and T2, and real segmentation labels. The network output is a segmentation map of each patient, including 4 types of segmentation labels: WT, TC, ET, and background area.

4.2. Evaluation Index. Four evaluation indicators are used in the experiment to evaluate the segmentation performance of the algorithm. The four evaluation indicators are Dice coefficient, Positive Predictive Value (PPV), Sensitivity, and Hausdorff distance. The calculation formulas of each indicator are as follows:

$$\text{Dice} = \frac{2TP}{FP + 2TP + FN},$$

$$\text{PPV} = \frac{TP}{FP + TP},$$

$$\text{Sensitivity} = \frac{TP}{TP + FN},$$

$$\text{Hans}(A, P) = \max \left[\sup_{a \in A} \inf_{p \in P} d(a, p), \sup_{p \in P} \inf_{a \in A} \text{dis}(a, p) \right], \quad (3)$$

where TP and TN represent pixels as True Positive and True Negative and FP and FN represent pixels as False Positive and False Negative. In the calculation formula of Hausdorff distance, sup represents the supremum, inf represents the inferior, A is the manually marked tumor area, a is the point on A , P is the predicted tumor area, p is the point on P , and $\text{dis}(a, p)$ represents the function used to calculate the distance between two points. Dice coefficient, PPV, and Sensitivity are used to evaluate the overlap between the real value and the predicted result. Large values of these three indicators represent good performance of the algorithm. Hausdorff distance is used to calculate the distance between the true value boundary and the predicted area boundary. A small Hausdorff value represents good performance of the algorithm.

4.3. Experimental Setup and Environment. Before starting to train the entire model, the hyperparameters in the model need to be set reasonably. This study uses a five-fold cross-validation method to determine the optimal hyperparameters to improve the effect and performance of the network. In the experiment, all weights are initialized randomly using $N(0, 1)$ Gaussian distribution. According to the graphics card and memory conditions, set batch_size to 1. The initial learning rate is set to 0.001 using Adam optimization algorithm [30]. The Adam optimization algorithm calculates the gradient's first-order moment estimation and second-order moment estimation to adapt the learning rate. The exponential decay rate of the first-order moment estimation is set to 0.88, and the exponential decay rate of the second-order moment estimation is set to 0.97. The comparison algorithms used in the experiment are CNN, U-net, and 3D U-net. The experimental environment is shown in Table 1.

4.4. Experimental Results and Analysis. The segmentation results of the four comparison models on the dataset are shown in Table 2. In order to compare the experimental data more vividly, this study uses a histogram to show the difference between each model in each indicator, as shown in Figure 6.

Table 2 shows the segmentation results of the four models for the three tissues. It can be seen from Table 2 that, in terms of Dice, PPV, and Sensitivity, 3D U-net and U-Net score higher than CNN. This shows that U-Net is more

TABLE 1: Experimental environment.

Configuration	Details
Operating system	Windows10 DircetX12 64
RAM	32G
Processor	Intel core i7-87003
Graphics card	Nvidia GeForce GTx 1080Ti
Programming framework	PyTorch
Architecture platform	CUDA8.0

TABLE 2: Image segmentation results of each comparison model.

Index	Segmented organization	CNN	U-net	3D U-net	Proposed
Dice	WT	0.7992	0.8285	0.8428	0.8787
	TC	0.7844	0.8006	0.8211	0.8456
	ET	0.7020	0.7602	0.7993	0.8004
	Mean	0.7619	0.7964	0.8211	0.8416
PPV	WT	0.8002	0.8473	0.8687	0.8992
	TC	0.7946	0.8331	0.8505	0.8668
	ET	0.7553	0.7717	0.7899	0.8321
	Mean	0.7837	0.8174	0.8364	0.866
Sensitivity	WT	0.8335	0.8586	0.8640	0.8899
	TC	0.8664	0.8903	0.8947	0.9078
	ET	0.7608	0.8192	0.8193	0.8335
	Mean	0.8202	0.856	0.8593	0.8771
Hausdorff	WT	2.9746	2.5762	2.6227	2.5002
	TC	1.8345	1.7481	1.7008	1.5101
	ET	3.2023	2.8350	2.8088	2.6443
	Mean	2.6704	2.3864	2.3774	2.2182

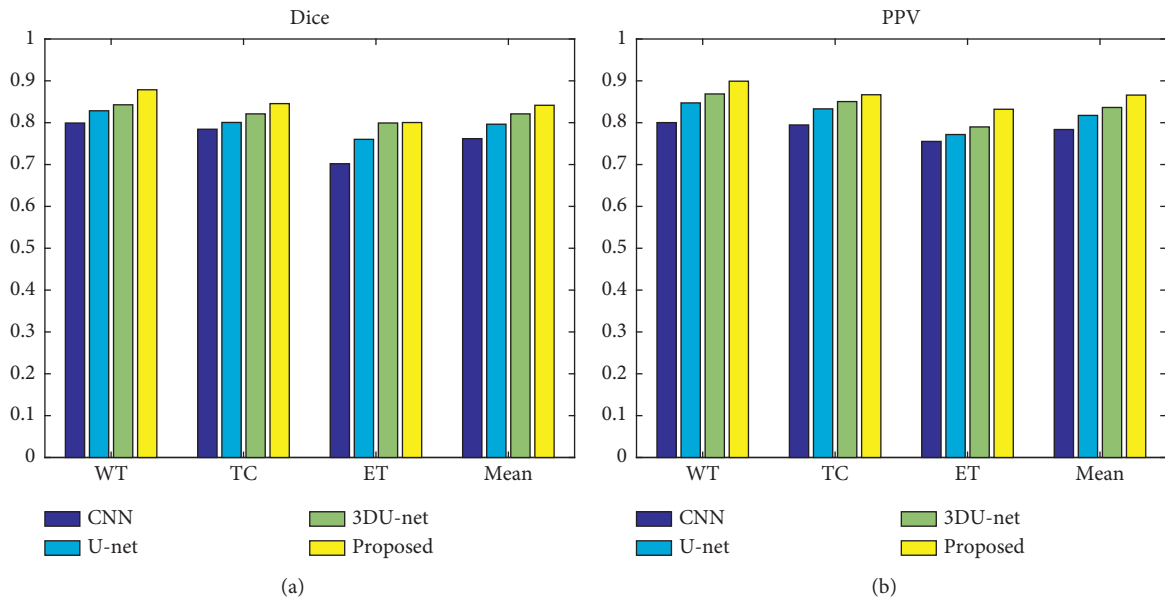


FIGURE 6: Continued.

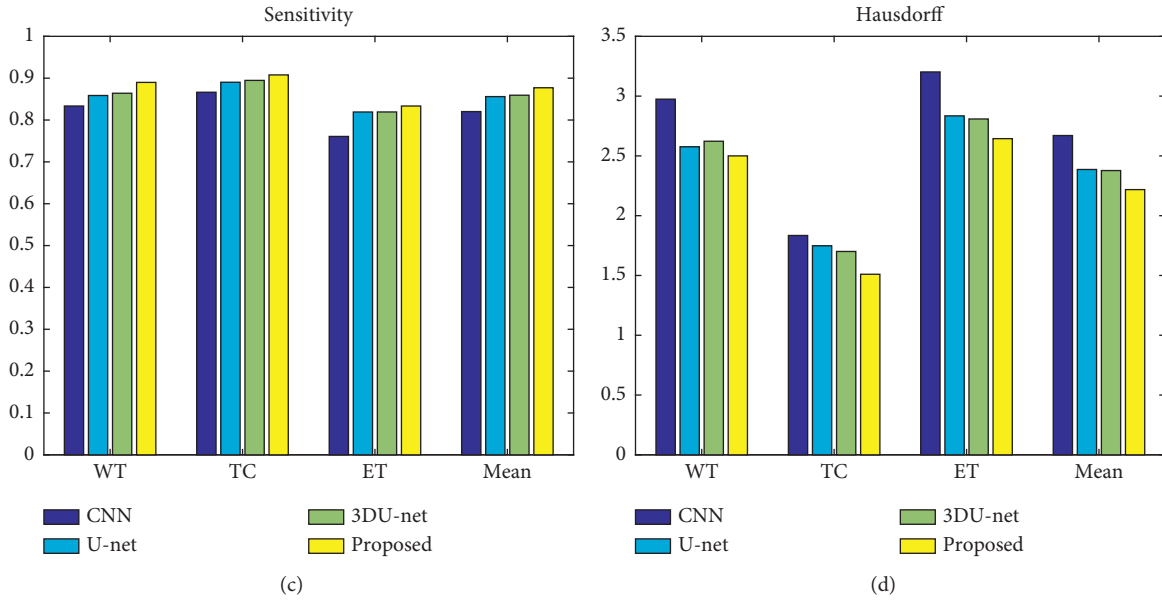


FIGURE 6: Segmentation comparison of each model on different indicators.

TABLE 3: The effect of different improved 3D U-net models on brain tumor segmentation.

Index	Segmented organization	[31]	[32]	[33]	Proposed
Dice	WT	0.7432	0.8328	0.8024	0.8787
	TC	0.6995	0.8006	0.7895	0.8456
	ET	0.6006	0.7670	0.7162	0.8004
	Mean	0.6811	0.8011	0.7964	0.8416
Hausdorff	WT	3.4675	2.8254	3.0052	2.5002
	TC	2.3006	1.8145	2.1723	1.5101
	ET	3.4783	2.8380	3.1016	2.6443
	Mean	3.0821	2.4926	2.7597	2.2182

suitable for medical IS. In the three indicators of Dice, PPV, and Sensitivity, 3D U-net has a larger index value than U-net. On the Hausdorff index, 3D U-net has a smaller value than U-net segmentation. This shows that 4 indicators all prove that 3D U-net has better segmentation effect than U-Net. From the data in the table, it can be seen that the method proposed in this paper has greatly improved the scores of Dice, PPV, and Sensitivity, and the value of Hausdorff scores has been significantly reduced. The 3D U-net that introduces the SECDC module can automatically learn the importance of different layers, thereby improving the segmentation accuracy and robustness of the model so that the segmentation effect of the model used in this article is optimal. Specifically, in terms of the Dice indicator, compared with CNN, U-Net, and 3D U-net, the model used has increased by 10.46%, 5.68%, and 2.50%, respectively. The Dice index of tumor segmentation of the whole tumor area exceeded 0.87, and the segmentation Dice index of tumor enhancement area reached 0.80. In the PPV index, compared with CNN, U-Net, and 3D U-net, the model used has increased by 10.50%, 5.95%, and 3.54%, respectively. In the Sensitivity index, compared with CNN, U-Net, and 3D U-net, the models used have increased by 6.94%, 2.46%, and 2.07%, respectively. In the Hausdorff index, compared with

CNN, U-Net, and 3D U-net, the models used are reduced by 16.93%, 7.05%, and 6.70%, respectively. From the analysis of the data, it can be seen that the model proposed in this paper is superior to other methods in terms of performance, accuracy, and sensitivity.

Since there are many improvement strategies for 3D U-net, many studies have applied their improved 3D U-net models to IS. In order to compare the effects of different improvement strategies on IS performance, and this paper compares the model used with other improved 3D U-Net models. Table 3 shows the performance comparison between the proposed model and other types of improved 3D U-Net models. In order to compare the experimental data more vividly, this study uses a histogram to show the difference between each model in each indicator, as shown in Figure 7.

Analyzing the segmentation performance of the three parts, the model segmentation in [31] has the worst effect, followed by [33], and then [32]. The IS performance based on the proposed model is relatively good. On the indicator Dice, for [31–33], the proposed model increased by 23.56%, 5.06%, and 5.68%, respectively. On the indicator Hausdorff, for [31–33], the proposed model increases by 28.03%, 11.01%, and 19.62%, respectively. According to the experimental

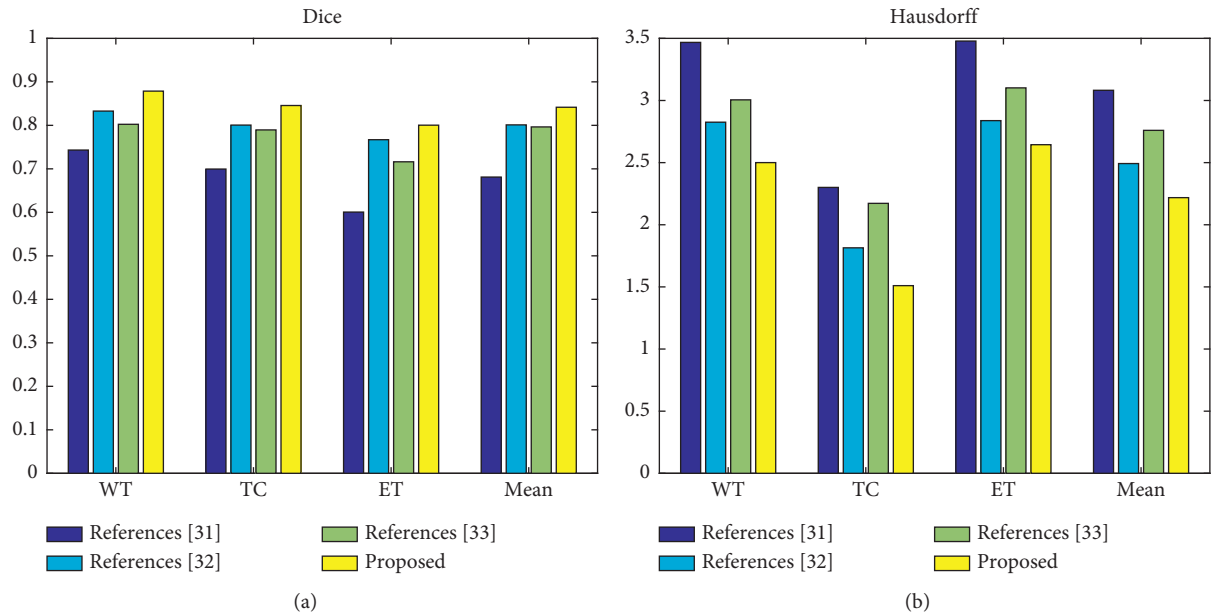


FIGURE 7: Segmentation comparison of various studies on different indicators.

results, it can be seen that the proposed model is superior to other methods in different segmentation areas of the tumor, and the overall segmentation performance is better.

5. Conclusion

Aiming at the unsatisfactory effect of traditional CNN on the segmentation of 3D medical images, this paper selects the U-Net model with higher performance for brain IS. Considering that the 3D U-Net model has problems such as complex network structure and large amount of calculation, this paper introduces the SECDC module into 3D U-Net, thereby constructing a high-precision lightweight segmentation model. The improved 3D U-Net network uses $1 \times 1 \times 1$ convolution to reduce the amount of parameters. Based on the normal convolution and the dilated convolution with the expansion rate of 2, the image features under different fields of view are effectively explored. At the same time, the 3D-SE module is introduced to effectively automatically learn the importance of different layers, thereby improving the robustness of the model. The experimental results on the BraTS2019 dataset prove the superiority of this method. However, in practical applications, there are still problems such as large sample labeling workload and long model segmentation time. According to these problems, the algorithm in this paper can be further optimized to achieve rapid network segmentation and efficient diagnosis by doctors. In addition, future work will continue to study the relationship between encoding and decoding and make full use of low-level features and semantic information to optimize the results.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Research Project of Jiangsu Province Colleges and Universities, under Grant 18KJD510011, and Jiangsu Province High-Level Key Professional Construction Project funding under Grant Su Jiaogao (2017) no. 17.

References

- [1] M. Gridach, "PyDiNet: pyramid dilated network for medical image segmentation," *Neural Networks*, vol. 140, pp. 274–281, 2021.
- [2] O. Avalos, E. Ayala, F. Wario, and M. Pérez-Cisneros, "An accurate Cluster chaotic optimization approach for digital medical image segmentation," *Neural Computing & Applications*, vol. 33, no. 16, pp. 10057–10091, 2021.
- [3] S. Rakesh and S. Mahesh, "Nodule segmentation of lung CT image for medical applications," *Global Transitions Proceedings*, vol. 2, no. 1, pp. 80–83, 2021.
- [4] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019.
- [5] L. Lu, M. Sun, Q. Lu, T. Wu, and B. Huang, "High energy X-ray radiation sensitive scintillating materials for medical imaging, cancer diagnosis and therapy," *Nanomaterials and Energy*, vol. 79, Article ID 105437, 2021.
- [6] D. Avola, L. Cinque, A. Fagioli, G. Foresti, and A. Mecca, "Ultrasound medical imaging techniques," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–38, 2021.
- [7] P. Li, T. Shi, Y. Zhao, and A. Lu, "Design of threshold segmentation method for quantum image," *International Journal of Theoretical Physics*, vol. 59, no. 2, pp. 514–538, 2020.

- [8] Y. Song, Y. L. Xu, and W. Gao, "A threshold segmentation method of chromosome microscope image," *Acta Microscopica*, vol. 28, no. 2, pp. 188–194, 2019.
- [9] Y. Yin, H. Luo, J. Sa, and Q. Zhang, "Study and application of improved level set method with prior graph cut in PCB image segmentation," *Circuit World*, vol. 46, no. 1, pp. 55–64, 2020.
- [10] M. Bernier, P.-M. Jodoin, O. Humbert, and A. Lalande, "Graph cut-based method for segmenting the left ventricle from MRI or echocardiographic images," *Computerized Medical Imaging and Graphics*, vol. 58, pp. 1–12, 2017.
- [11] S. Liu, Z. Liu, and K. Jiang, "Image target segmentation method based on fuzzy Renyi entropy and region growing," *Systems Engineering and Electronics*, vol. 40, no. 8, pp. 1693–1701, 2018.
- [12] J. Ma, K. Du, L. Zhang, F. Zheng, J. Chu, and Z. Sun, "A segmentation method for greenhouse vegetable foliar disease spots images using color information and region growing," *Computers and Electronics in Agriculture*, vol. 142, pp. 110–117, 2017.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Neural Information Processing Systems*, vol. 25, no. 2, pp. 1097–1105, 2012.
- [14] P. Matlani and M. Shrivastava, "Hybrid deep VGG-NET convolutional classifier for video smoke detection," *Computer Modeling in Engineering and Sciences*, vol. 119, no. 3, pp. 427–458, 2019.
- [15] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 3431–3440, 2015.
- [16] O. Ronneberger, P. Fischer, T. Brox et al., "U-net: convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer Assisted Intervention*, Springer, New York, NY, USA, pp. 234–241, 2015.
- [17] V. Iglovikov and A. Shvets, "TernausNet: U-net with VGG11 encoder pre-trained on ImageNet for image segmentation," 2018, <https://arxiv.org/abs/1801.05746>.
- [18] O. Oktay, J. Schlemper, L. L. Folgoc et al., "Attention U-net: learning where to look for the pancreas," 2018, <https://arxiv.org/abs/1804.03999>.
- [19] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: redesigning skip connections to exploit multiscale features in image segmentation," *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 1856–1867, 2020.
- [20] H. Sun, C. Li, B. Liu et al., "AUNet: attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms," *Physics in Medicine and Biology*, vol. 65, no. 5, Article ID 055005, 2020.
- [21] F. Milletari, N. Navab, and S. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the International Conference on 3D Vision*, pp. 565–571, Stanford, CA, USA, October 2016.
- [22] T. G. W. Boers, Y. Hu, E. Gibson et al., "Interactive 3D U-net for the segmentation of the pancreas in computed tomography scans," *Physics in Medicine and Biology*, vol. 65, no. 6, Article ID 065002, 2020.
- [23] X. Li, H. Chen, X. Qi, Q. Du, C. W. Fu, and P. A. Heng, "H-DenseUNet: hybrid densely connected UNet for ltumor segmentation from CT volumes," *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
- [24] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-net: learning dense volumetric segmentation from sparse annotation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424–432, Springer, Athens, Greece, October 2016.
- [25] Y. Wang, L. Zhao, M. Wang, and Z. Song, "Organ at risk segmentation in head and neck CT images using a two-stage segmentation framework based on 3D U-net," *IEEE Access*, vol. 7, pp. 144591–144602, 2019.
- [26] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 2011–2023, 2020.
- [27] B. Li, Z.-W. Tan, P. P. Shum, C. Wang, Y. Zheng, and L. j. Wong, "Dilated convolutional neural networks for fiber Bragg grating signal demodulation," *Optics Express*, vol. 29, no. 5, pp. 7110–7123, 2021.
- [28] Y. Wang, S. Hu, G. Wang, C. Chen, and Z. Pan, "Multi-scale dilated convolution of convolutional neural network for crowd counting," *Multimedia Tools and Applications*, vol. 79, pp. 1057–1073, 2020.
- [29] M. Havaei, A. Davy, D. Warde-Farley et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.
- [30] A. Barakat and P. Bianchi, "Convergence and dynamical behavior of the ADAM algorithm for nonconvex stochastic optimization," *SIAM Journal on Optimization*, vol. 31, no. 1, pp. 244–274, 2021.
- [31] M. Bhalerao and S. Thakur, "Brain tumor segmentation based on 3D residual U-Net," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11993, pp. 218–225, 2020.
- [32] N. Boutry, J. Chazalon, E. Puybureau, G. Tochon, H. Talbot, and T. Géraud, "Using separated inputs for multimodal brain tumor segmentation with 3d u-net-like architectures," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11992, pp. 187–199, 2020.
- [33] W. Jun, X. Haoxiang, and Z. Wang, "Brain tumor segmentation using dual-path attention U-net in 3D MRI images," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12658, pp. 183–193, 2021.

Research Article

Analysis of Effectiveness and Performance Prediction of Sports Flipped Classroom Teaching Based on Neural Networks

Wei Xu , Wenying Xiong, Zhe Shao, and Yun Li 

College of Physical Education and Health, Jiangxi University of Traditional Chinese Medicine, Nanchang 330000, Jiangxi, China

Correspondence should be addressed to Yun Li; 20142031@jxutcm.edu.cn

Received 28 April 2021; Revised 22 June 2021; Accepted 4 July 2021; Published 31 July 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Wei Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traditional physical education methods are unable to meet this requirement due to the practical nature of sports skill teaching. As a result, as the times demanded, the flipped classroom based on neural network technology arose. It has the potential to not only promote the modernization of physical education but also to ensure that it has a positive educational impact. This is a mode of instruction. Furthermore, colleges and universities are increasingly focusing on college students' overall quality development. A method for predicting college students' sports performance using a particle swarm optimization neural network is proposed to accurately predict sports performance and provide a reliable analysis basis for the establishment of sports teaching goals. Neural networks are used in the model. The particle swarm optimization algorithm optimizes the variance and weights of the neural network to improve the accuracy of college students' sports performance predicted by the neural network by updating the particle position and speed through the two extreme values of individual extreme values and global extreme values. Teachers always play the role of the facilitator and helper in the teaching process, which realizes the transformation of teachers' and students' self-positioning, allows students to better play the lead role, and stimulates students' interest in learning.

1. Introduction

Flipped classroom [1–3] is a product of modern development, and internet technology is an important support for its implementation: teachers send teaching videos [4–6] made by themselves or by other teachers on the internet to students via the learning platform [7] so that students can complete preclass learning tasks under the guidance of teachers in the classroom. The flipped classroom is popular in theoretical subjects but is rarely used in physical education [8], which is a practical course. However, the emergence of this teaching model allows for the combination of the internet and the physical education classroom. The flipped classroom is a teaching model that reinvents the traditional classroom process of teaching sports skills, bringing the connection between before, during, and after class closer together, and extending the actual teaching and learning time between teachers and students. PE teachers must create small videos based on the teaching content of each lesson in this mode of instruction. Small video

produced should be concise and complete the explanation of a lesson's content within 5 to 10 minutes. The level of each physical education teacher varies, as does the quality of the short video produced.

With the ongoing development and popularization of internet technology [9], students are increasingly relying on the convenience that the internet provides. The introduction of internet-based flipped classrooms will undoubtedly cause a revolution in the educational arena. Both students' learning and teachers' teaching have changed as a result of this revolution. Students' learning processes are more independent, and teachers are better able to teach students based on their aptitude. Individuality is a goal for college students, who want to think for themselves. Following the clarification of the learning objectives, they make full use of their ability to consult and gain a preliminary understanding of materials on their own. They understand and learn sports skills before class through sports online courses and then internalize and apply what they have learned in class to improve learning efficiency [10].

Students' mastery of sports skills in physical education cannot be truly realized solely through online video learning. Learning sports skills is a unique learning activity that necessitates physical participation and is inextricably linked to classroom exercises. The flipped classroom teaching model allows students to finish their knowledge learning before class. Classroom is a place for teachers to answer students, to interact with teachers and students, and to apply knowledge. The popularization of the internet makes this teaching model more feasible. The flipped classroom based on the neural network [11–14] will extend students' learning time and learning space and is no longer limited to the fixed weekly class time and place. The traditional PE teaching classroom is flipped to a certain extent, and internet is used for learning before and after class. This teaching mode [15] can organically combine the learning of sports skills with online resources under the internet and improve students' enthusiasm for sports learning to a certain extent. Furthermore, students' performance predictions in the sports flipped classroom have a significant impact on classroom interaction. As a result, based on particle swarm optimization neural networks [16–18], this paper proposes a method for predicting college students' sports performance [19].

Machine learning [20] is the science of studying how to use computers to simulate or realize human learning activities. It is one of the most intelligent and cutting-edge research fields in artificial intelligence. Since the 1980s, machine learning, as a way to realize artificial intelligence [21–25], has aroused widespread interest in the artificial intelligence community. Especially in the past ten years, the research work in the field of machine learning has developed rapidly, and it has become an important part of artificial intelligence. Machine learning is not only used in knowledge-based systems but also widely used in many fields [26–30] such as natural language understanding, non-monotonic reasoning, machine vision [31–33], and pattern recognition [31, 32]. Therefore, it is feasible to use machine learning to predict the effect and performance of sports flipped classroom teaching.

Following are the main innovation points of this paper:

- (1) This paper investigates the use of neural network technology in the sports flip classroom to improve communication between teachers and students, strengthen interaction and cooperation between the two, provide timely and comprehensive sports skill learning for students, improve students' sports skills, and stimulate students' interest in learning sports skills
- (2) A prediction method for college students' sports performance based on the particle swarm optimization neural network is proposed to achieve high-efficiency and high-precision prediction of college students' sports performance in order to improve the accuracy of college students' sports performance prediction

2. Related Work

The term “flipped classroom” refers to rearranging time inside and outside of the classroom in order to transfer learning decision-making power from the teacher to the student. Students can better display their subjective initiative in this classroom teaching model, and teachers can better play their guiding role. Before class, students study independently by watching video lectures and chatting with classmates via the internet. Communication and exchanges between students and teachers have increased as a result of the internet. Students can learn a variety of online courses outside of the classroom if they have access to the internet. The internet era gave birth to this flipped teaching classroom. It has disrupted the traditional teaching classroom structure and resulted in a slew of educational changes.

Motor skills [34] refer to the human body in the movement to master and effectively complete a special movement of an ability. Some scholars believe that it consists of two parts: one is the description of the rules for performing the action, that is, the procedure of the action, and the other is the actual muscle movement, which gradually becomes precise and coherent through practice and feedback. Physical education is a teaching process that takes physical activity as the means, while the teaching of sports skills needs to improve students' ability of sports skills through physical education. Students must feel the joy of the environment, democratic and harmonious interpersonal relationships, individuality liberation, and the learning environment in order to learn sports skills.

Al Zahrani [35] elaborated on the impact of flipped classrooms on the creative thinking of higher education students. In the study, they recruited students from the School of Education of King Abdulaziz University in Saudi Arabia and conducted group teaching experiments on them. Their research results showed that flipped classrooms can improve students' creativity, especially in knowledge, and master the fluency, flexibility, and novelty aspects. Sajid et al. [36] discussed in their article that flipped classroom student-centered active learning replaces passive learning. This learning method improves critical thinking and application skills, including information retention capacity. Larsson and Nyberg [37] stated that sport is the key to physical education. Schools in Sweden are attempting to incorporate social constructivist knowledge and learning concepts into physical education, with an emphasis on students' motor skills. They discussed how to intervene in students' behavior and how to communicate between students and teachers in the article. They believed that students must participate in the teaching process in order for teachers to gain a better understanding of their students and provide better guidance and assistance in the development of motor skills.

3. Methodology

3.1. Particle Swarm Neural Network Model

3.1.1. Neuron Model. In 1943, McCulloch and Pitts [38] used a simple model to express the neuron structure and created an artificial neuron model, which is now known as the “M-P neuron model,” as shown in Figure 1.

The following is the calculation equation for the neuron’s output in Figure 1:

$$y = f\left(\sum_{i=1}^n w_i x_i - \theta\right), \quad (1)$$

where θ is the activation threshold of the neuron we mentioned earlier and the function $f(\cdot)$ is also called the activation function. As shown in the figure, the function $f(\cdot)$ can be expressed by a step equation, which is activated above the threshold; otherwise, it is inhibited, but this is a bit too crude because the step function is not smooth, discontinuous, and nonderivable, so our more commonly used method is to use the sigmoid function to represent the function $f(\cdot)$.

Figure 2 shows the sigmoid activation function, and its calculation equation is as follows:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2)$$

3.1.2. Particle Swarm Optimization. A particle swarm optimization method to solve the optimization problem is summarized by observing the behavior of birds looking for food. It is necessary to search for particles in the optimized space in order to discover the essence of the optimization problem. The speed of each particle determines the distance and direction of the bird’s flight. Every particle has a fitness value, and the fitness value is determined by the optimization. The purpose of all particles in space is to find the optimal particle action. The initial random solution (random particle) of the particle swarm optimization method is continuously iterated to find the optimal solution. The particle “tracks” the two extreme values in each iteration to ensure that the particle itself can be updated. These two extreme values are t_{best} and u_{best} , respectively. t_{best} is the optimal solution found by the particle itself, called the individual extreme value, and u_{best} is the optimal solution found by the entire population, called the global extreme value. The particle position and velocity can be updated by the following equation:

$$\begin{aligned} w &= w + x_1 \times y_1 \times (t_{\text{best}} - \text{pre}) + x_2 \times y_2 \times (u_{\text{best}} - \text{pre}), \\ \text{pre} &= \text{pre} + w. \end{aligned} \quad (3)$$

The velocity of the particle is represented by w , the position of the particle itself is represented by pre , the random numbers between (0, 1) are y_1 and y_2 , and the learning factors are x_1 and x_2 , respectively.

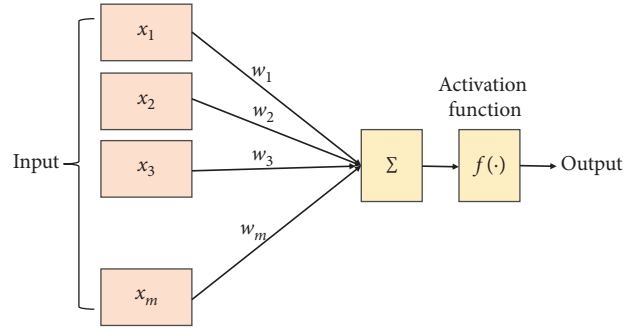


FIGURE 1: M-P neuron model.

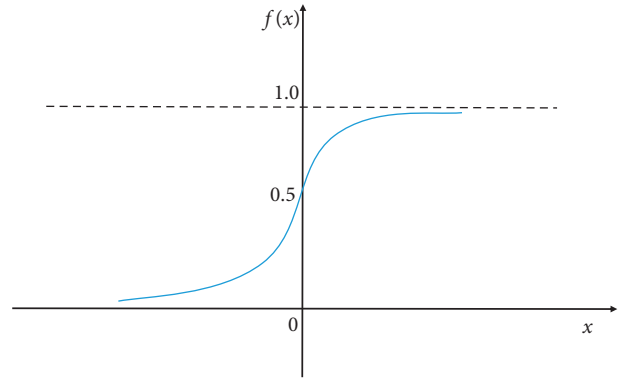


FIGURE 2: Sigmoid function.

Because the first particle velocity w on the right side of equation (3) is random and has no memory, new regions will be explored during the search, demonstrating strong global optimization power. However, in practical applications, it is necessary to conduct a global search first to improve the search’s convergence speed and then to conduct a local search to improve the accuracy of the solution obtained. The optimized calculation equation is as follows:

$$w = \sigma \times w + x_1 + y_1 \times (t_{\text{best}} - \text{pre}) + x_2 \times y_2 \times (u_{\text{best}} - \text{pre}). \quad (4)$$

The inertia weight is expressed as σ . If the value of σ is large, the PSO method has a strong global search ability; otherwise, it has a strong local search ability.

3.1.3. Improve Variance and Weights. When using the particle swarm optimization method to optimize the neural network, it is necessary to optimize the variance σ_i ($i = 1, 2, \dots, h$) and the weight v_0, v_1, \dots, v_k of the neural network basis function, the dimensions of the two parameters are related to the network structure, and the weight and variance of the neural network are uniformly coded. A set of neural network weights and variances are described by a particle. The root mean square error (RMSE) can be used as a fitness function to reflect the particle’s approximation error. The following is the calculation equation for the fitness value of the i th particle:

$$\text{RMSE}(i) = \sqrt{\frac{1}{a} \sum_{j=1}^a (y_1^j - y_2^j)^2}, \quad (5)$$

where a is the total number of samples and the true output value and neural network prediction value are described by y_1^i and y_2^i , respectively.

3.2. Effectiveness Analysis and Performance Prediction of the Sports Flipped Classroom. College students' sports performance is modeled using the particle swarm optimization neural network (PSO) model. In terms of time order, college students' physical education achievements in the past are typical. For the modeling of college students' physical education achievements, a particle swarm optimization neural network model is used. Assume that the initial time series of college students' physical education achievements is shown in the following equation:

$$A^{(0)}(p) = \{a^{(0)}(1), a^{(0)}(2), \dots, a^{(0)}(m)\}. \quad (6)$$

Operate equation (6) with the first-order differential; then, we can get

$$A^{(1)}(p) = \{a^{(1)}(1), a^{(1)}(2), \dots, a^{(1)}(m)\}. \quad (7)$$

The first-order mean operation equation (7) can obtain the first-order mean value sequence as follows:

$$B^{(1)}(p) = \{b^{(1)}(1), b^{(1)}(2), \dots, b^{(1)}(m)\}. \quad (8)$$

Then, a predictive model is established, and the calculation equation is as follows:

$$\frac{fa^{(1)}}{fp} + ca^{(1)} = \lambda. \quad (9)$$

Solving the above equations, the whitening differential equation that can predict the performance of the sports flipped classroom is as follows:

$$\hat{a}^{(1)}(h+1) = \left(a^{(0)}(1) - \frac{\lambda}{c} \right) d^{-ch} + \frac{\lambda}{c}. \quad (10)$$

Next, the neural network parameters are obtained by training according to the input and output sample data, and the optimal target mapping of equation (10) is completed.

4. Experiments and Results

4.1. Experimental Setup. The subjects of the experiment are the two male classes of the 2018 physical education training major of the School of Sports Science of Qufu Normal University—the general volleyball class (17 people). Before determining the experimental objects, preexperiment interviews and physical fitness tests are conducted for the two classes of students. There are significant differences, students who have a foundation in volleyball and other students in the class also participate in the learning, and the data obtained from the learning results are not included in the experiment.

The final students who actually participated in the experiment are listed and determined. Both the experimental class and the control class were of 16 students. The students participating in the teaching experiment are all second-year university students. When students are exposed to new things, they have a strong curiosity about them, as well as strong learning and accepting abilities. Students use mobile phones, computers, and iPads for entertainment, learning, and other purposes in their daily lives. As a result, there will be no problem with them using the school's online program, and there will be no shortage of smart devices and technology.

From March 4 to June 21, a teaching experiment was conducted. Professional teachers provided guidance throughout the process, which included discussing teaching progress with the teacher and carrying out the teaching experiment with the teacher's consent. In the teaching experiment, both the experimental class and the control class are taught twice a week. Each class is two hours. One of the two hours is used to complete experimental teaching, and the remaining one is used to complete the school's teaching plan. The total class hours for teaching experiments are 32 hours.

4.2. Experimental Results. Start at midpoint A of the end line, and the stopwatch starts at the moment of start. Run along line 1 to intersection point B of the end line and the sideline, touch the volleyball, and then quickly turn back to midpoint A of the end line. Touch the volleyball, and then quickly run along line 2 to intersection point C of the offensive line and the sideline, touch the volleyball, and then quickly follow 2. Fold back to midpoint A of the end line, touch the volleyball, and then quickly run to midpoint D of the offensive line along the 3rd line. Touch the volleyball, continue to fold back to midpoint A of the end line along the 3rd line, and then run along the 4th line towards the offensive line. Touch the volleyball at intersection point E of the sideline, and then fold back to midpoint A of the end line along line 4. Touch the volleyball, finally run along line 5 to intersection point F between the end line and the sideline, fold back to point A , complete the half-meter run test, and stop timing. At six points A , B , C , D , E , and F , arrange a student to place a volleyball, and supervise the tester to touch the volleyball during the movement. The moving route is shown in Figure 3. In addition, Figure 4 also gives the experimental scene of the physical fitness test.

Table 1 shows that the neural network-based method used in this article outperforms the traditional method significantly. BP network [33] is slightly inferior to the other two neural network methods. As a result, the algorithm presented in this article is capable of analyzing and predicting the effectiveness of sports flipped classroom teaching.

It can be seen from Figure 5 that when the sports flipped classroom performance predicted by this method is used, the sports performance prediction accuracy of each sports item is higher than 9500, while the accuracy of the other two methods is below 90% for both, indicating that the sports

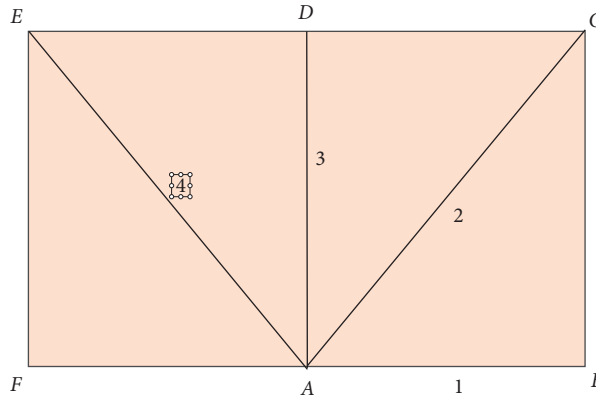


FIGURE 3: Experimental site.



FIGURE 4: Physical fitness test.

TABLE 1: Experimental results of the speed test.

Method	N	Mean	SD	Acc
Traditional	16	18.11	0.77	0.81
BP	16	18.12	0.81	0.85
Hinojo-Lucena et al.	16	18.62	0.94	0.92
Ours	16	18.63	0.96	0.92

flipping classroom performance prediction of the method in this paper has strong versatility and high precision and can be widely used in the future study of college students' sports performance.

4.3. *Ablation Experiment.* We conducted an ablation experiment. We input 3, 5, and 8 factors to observe their influence on the experimental results. The experimental results are shown in Table 2.

It can be clearly seen from Table 2 that, with the increase of factors, the model's prediction accuracy rate increases, but at the same time, we found that, from 5 factors to 8 factors, the accuracy rate has not greatly improved, considering the collection of data, and preprocessing can be a tedious work; therefore, this article believes that it is more appropriate to consider five factors.

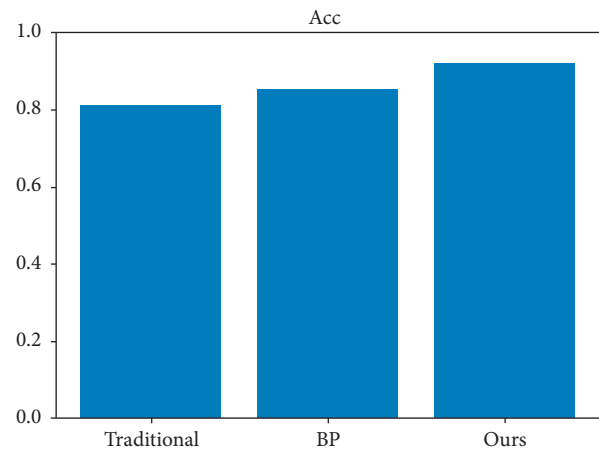


FIGURE 5: Experimental results of accuracy comparison.

TABLE 2: Results of ablation experiments.

Input	Acc
3 factors	0.8636
5 factors	0.9207
8 factors	0.9258

5. Conclusion

To accurately predict sports performance and provide a reliable analysis basis for the establishment of sports teaching goals, a method for predicting college students' sports performance based on the particle swarm optimization neural network is proposed. The model is based on neural networks. The particle swarm optimization algorithm updates the particle position and speed through the two extreme values of individual extreme values and global extreme values to optimize the variance and weights of the neural network to enhance the accuracy of college students' sports performance predicted by the neural network. In the teaching process, teachers always take the role of the facilitator and helper to carry out teaching, which realizes the transformation of teachers' and students' self-positioning, can better play the main role of students, and stimulate students' interest in learning.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

All the authors declare no conflicts of interest.

References

- [1] G. Akçayır and M. Akçayır, "The flipped classroom: a review of its advantages and challenges," *Computers & Education*, vol. 126, pp. 334–345, 2018.
- [2] K. F. Hew and C. K. Lo, "Flipped classroom improves student learning in health professions education: a meta-analysis," *BMC Medical Education*, vol. 18, no. 1, pp. 38–12, 2018.
- [3] P. Strelan, A. Osborn, and E. Palmer, "The flipped classroom: a meta-analysis of effects on student performance across disciplines and education levels," *Educational Research Review*, vol. 30, Article ID 100314, 2020.
- [4] S. L. Cresswell, W. A. Loughlin, M. J. Coster, and D. M. Green, "Development and production of interactive videos for teaching chemical techniques during laboratory sessions," *Journal of Chemical Education*, vol. 96, 2019.
- [5] A. Meseguer-Martinez, A. Ros-Galvez, and A. Rosa-Garcia, "Satisfaction with online teaching videos: a quantitative approach," *Innovations in Education & Teaching International*, vol. 54, no. 1, pp. 62–67, 2017.
- [6] P. Born, N. P. Nguyen, R. Grambow, D. Meffert, and T. Vogt, "Embedding tennis-specific teaching videos into long-term educational concepts to improve movement learning and technique performances," *Journal of Physical Education and Sport*, vol. 18, no. 1, pp. 255–261, 2018.
- [7] N. F. B. Casillano, "Challenges of implementing an E-learning platform in an Internet struggling province in the Philippines," *Indian Journal of Science and Technology*, vol. 12, no. 10, pp. 1–4, 2019.
- [8] A. Casey and A. MacPhail, "Adopting a models-based approach to teaching physical education," *Physical Education and Sport Pedagogy*, vol. 23, no. 3, pp. 294–310, 2018.
- [9] A. Zohrabi, J. Karimi, and A. Mohebi, "Explaining the applications of Internet of things in Iran's sports industry: a sustainable development approach," *Communication Management in Sport Media*, vol. 7, no. 3, pp. 91–99, 2020.
- [10] T. Tang, A. M. Abuhmaid, M. Olaimat, D. M. Oudat, M. Aldhaeabi, and E. Bamanger, "Efficiency of flipped classroom with online-based teaching under COVID-19," *Interactive Learning Environments*, pp. 1–12, 2020, In press.
- [11] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local-global multiple correlation filters for robust visual tracking with kalman filter redetection," *Sensors*, vol. 21, no. 4, Article ID 1129, 2021.
- [12] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, 2020, In press.
- [13] X. Ning, X. Wang, and S Xu, "A review of research on co-training. concurrency and computation: practice and experience," *Concurrency and Computation Practice and Experience*, 2021, In press.
- [14] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, pp. 1–22, 2021.
- [15] Q. Luo, "Research on the teaching mode of physical education in colleges and universities," in *Proceedings of the 2017 2nd International Conference on Education, Sports, Arts and Management Engineering (ICESAME 2017)*, pp. 716–719, Zhengzhou, China, June 2017.
- [16] Y. Zhang, W. Li, L. Zhang, X. Ning, L. Sun, and Y. Lu, "AGCNN: adaptive gabor convolutional neural networks with receptive fields for vein biometric recognition," *Concurrency and Computation: Practice and Experience*, Article ID e5697, 2020, In press.
- [17] W. Cai, Y. Song, and Z. Wei, "Multimodal data guided spatial feature fusion and grouping strategy for e-commerce commodity demand forecasting," *Mobile Information Systems*, vol. 2021, Article ID 5568208, 14 pages, 2021.
- [18] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.
- [19] F. Muñoz-Bullón, M. J. Sanchez-Bueno, and A. Vos-Saz, "The influence of sports participation on academic performance among students in higher education," *Sport Management Review*, vol. 20, no. 4, pp. 365–378, 2017.
- [20] C. Yan, G. Pang, X. Bai et al., "Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2021, In press.
- [21] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.
- [22] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.
- [23] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person Re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, In press.
- [24] J. Zhang, J. Sun, J. Wang, and X. G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–14, 2020, In press.

- [25] X. Ning, P. Duan, W. Li, and S. Zhang, "Real-time 3D face alignment using an encoder-decoder network with an efficient deconvolution layer," *IEEE Signal Processing Letters*, vol. 27, pp. 1944–1948, 2020.
- [26] J. Zhang, X. Jin, J. Sun, J. Wang, and K. Li, "Dual model learning combined with multiple feature selection for accurate visual tracking," *IEEE Access*, vol. 7, pp. 43956–43969, 2019.
- [27] X. Ning, K. Gong, W. Li, and L. Zhang, "JWSAA: joint weak saliency and attention aware for person re-identification," *Neurocomputing*, vol. 453, pp. 801–811, 2020.
- [28] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.
- [29] Y. Jiang, X. Gu, D. Wu et al., "A novel negative-transfer-resistant fuzzy clustering model with a shared cross-domain transfer latent space and its application to brain CT image segmentation," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 1, pp. 40–52, 2020.
- [30] Y. Gu, A. Chen, X. Zhang, C. Fan, K. Li, and J. Shen, "Deep learning based cell classification in imaging flow cytometer," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 2, pp. 18–27, 2021.
- [31] Y. Jiang, K. Zhao, and K. Xia, "A novel distributed multitask fuzzy clustering algorithm for automatic MR brain image segmentation," *Journal of Medical Systems*, vol. 43, no. 5, pp. 118–119, 2019.
- [32] Z. Huang, P. Zhang, R. Liu, and D. Li, "Immature apple detection method based on improved Yolov3," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 9–13, 2021.
- [33] L. Huang, G. Xie, W. Zhao, Y. Gu, and Y. Huang, "Regional logistics demand forecasting: a bp neural network approach," *Complex & Intelligent Systems*, pp. 1–16, 2021, In press.
- [34] S. W. Logan, S. M. Ross, K. Chee, D. F. Stodden, and L. E. Robinson, "Fundamental motor skills: a systematic review of terminology," *Journal of Sports Sciences*, vol. 36, no. 7, pp. 781–796, 2018.
- [35] A. M. Al-Zahrani, "From passive to active: the impact of the flipped classroom through social learning platforms on higher education students' creative thinking," *British Journal of Educational Technology*, vol. 46, no. 6, pp. 1133–1148, 2015.
- [36] M. R. Sajid, A. F. Laheji, F. Abothenain, Y. Salam, D. AlJayar, and A. Obeidat, "Can blended learning and the flipped classroom improve student learning and satisfaction in Saudi Arabia?" *International Journal of Medical Education*, vol. 7, pp. 281–285, 2016.
- [37] H. Larsson and G. Nyberg, "'It doesn't matter how they move really, as long as they move.' Physical education teachers on developing their students' movement capabilities," *Physical Education and Sport Pedagogy*, vol. 22, no. 2, pp. 137–149, 2017.
- [38] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

Research Article

A New Embedded Estimation Model for Soil Temperature Prediction

Xuezhi Wang ^{1,2}, Wenhui Li ^{1,2} and Qingliang Li ^{2,3}

¹College of Computer Science and Technology, Jilin University, Changchun 130012, China

²Symbol Computation and Knowledge Engineer of Ministry of Education, Jilin University, Changchun, Jilin Province, China

³School of Computer Science and Technology, Changchun Normal University, Changchun 130032, China

Correspondence should be addressed to Qingliang Li; liqingliang@ccsfu.edu.cn

Received 5 June 2021; Accepted 8 July 2021; Published 17 July 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Xuezhi Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous development of Earth science, soil temperature has received more and more attention in Earth system research as an important parameter. The change of soil temperature (T_s) in different regions and related time series is affected by many factors, which bring certain difficulties to the accuracy of soil temperature prediction and the robustness of the algorithm. In this paper, an embedded network prediction model based on the gated recurrent unit (GRU) model is proposed to learn the local and global features of historical temperature for improving the prediction performance of soil temperature. We input different steps into the GRU model, and the output is weighted to obtain the final prediction result. In order to obtain the global characteristics of soil temperature, we connect the previous steps to the output layer directly, and the local characteristics of soil temperature are obtained through the following steps. This paper uses the soil temperature data from two meteorological stations (Laegern and Fluehli) in Switzerland as the input data to predict the soil temperature for different soil depths (5 cm, 10 cm, and 15 cm) at different time points (6 hrs, 12 hrs, and 24 hrs), using RMSE, MAE, MSE, and R^2 performance indicators as evaluation criteria to verify the accuracy of prediction. As the experimental results show, our method has the best performance compared to the others (artificial neural networks (ANN), extreme learning machine model (ELM), long short-term memory network (LSTM), gated recurrent unit network (GRU)). In particular, we estimated the soil temperature at the soil depth of 10 cm of the Fluehli station in the coming 6 hrs; our method achieved the best performance; and, meanwhile, our model achieved the maximum value of R^2 (0.9914) and the minimum values of RMSE (0.4668), MAE (0.2585), and MSE (0.2214) compared with the other four models. Therefore, our model can not only predict the soil temperature at different depths but also improve the accuracy.

1. Introduction

Geoscience has played an important role in social development and economic construction; soil temperature (T_s) and its daily fluctuations are among the most vital meteorological parameters in Earth sciences, such as agriculture, forestry, and geology [1, 2]; and it is an important variable of land-atmosphere interactions [3]. Meanwhile, there are many elements that affect the change of soil temperature; for example, the change of soil depth has a prominent effect on soil temperature. Research has shown that, in the processing of plant growth, shallow soil has a significant impact on seed germination, while deep soil affects root absorption activity [4]. Therefore, the accurate prediction of soil temperature at

different depths can be used to guide practical applications in some fields, which is instead of using traditional sensors manually for on-site measurement [5].

Currently, most of the soil temperature prediction methods use environmental factors to estimate [6]. However, the data collected in some regions is unavailable that cannot be used to predict, which will reduce the accuracy of model predictions [7]. Therefore, this paper recommends using time series as the input data for the soil temperature prediction model.

In recent years, researchers usually use methods based on physical models to predict soil temperature through the heat transfer mechanism of the soil itself mainly [6, 8]. However, there are many limitations in practical applications due to the

physical model parameterization and scale issues [9]. With the continuous development of the machine learning method, it has been widely used in Earth sciences [10–12]. Ghorbani et al. proposed a method based on the support vector machine to estimate the soil field capacity and permanent wilting point [13]. And it also plays an important role in the soil temperature field [14]. The extreme learning machine is used to predict the soil temperature for improving the accuracy by Feng et al. [15]. Furthermore, LSTM has also received attention from researchers [16].

For the machine learning method, the artificial neural network (ANN) [2, 17, 18] and ELM [15, 19] can learn the features from the input data without using a physical model, so there is no need to understand their internal physical processes. The artificial neural network has strong self-adaptation and self-learning capabilities and can continuously update the parameters in the model to make the output value closer to the real value, so it is used as a soil temperature prediction model usually [20–22]. Bilgili proposed a method based on the artificial neural network to predict monthly average soil temperature [23]. Mehdizadeh et al. used the model based on feedforward backpropagation neural networks (FFBPNN) and gene expression programming (GEP) to estimate the daily soil temperature at different depths [24]. When the traditional artificial neural network is used to predict the soil temperature, the accuracy of the output results is reduced because the correlation of the time series is not considered. For solving this problem, many researchers merge genetic algorithms into artificial neural networks to optimize neural networks [20, 25, 26]. However, the genetic algorithm is not efficient and prone to premature convergence. Therefore, this method still needs further research.

Deep learning methods are widely used to deal with time series data. Compared with recurrent neural networks (RNN) and LSTM, GRU has a simpler structure and can solve the problem of long-term dependence [27]. Therefore, this paper chooses the model based on the GRU network to predict soil temperature. The model uses hidden states to convey the information and process the relevance of time series data. It is widely used in many fields due to its special network structure. Liu et al. proposed a method, GRU-based nonlinear predictive denoising autoencoders for fault diagnosis of rolling bearing [28]. Miao and Hung designed the Conv-GRU model to estimate the water level, and the results show the effectiveness of the method [29]. Rui et al. combined GRU and LSTM models to predict traffic flow [30]. According to our research, the GRU network had not been used for soil temperature prediction.

The following questions are the focus of this article. The first one is how to choose the input data to T_s estimation model. The estimation of T_s is affected by the past T_s . Although the relevant meteorological data have some impact on T_s estimation, the accuracy of the model for T_s estimation will be affected by the errors between the provided data and the real data. Consequently, this paper concentrates on the time series of data. The other one is about the network model construction of the method in our paper. In the GRU network, the information is transferred by updating the cell state and the parameters in the hidden state. As the steps of

the time series increase, the correlation between the initial data and the output data will be decreased, which will lead to a decrease in prediction accuracy.

The motivation of this paper is to solve the problem of long-term serial dependence of soil temperature data, which leads to a decrease in prediction accuracy. With the goal, this paper proposed a new embedded estimation model based on the GRU network for T_s estimation. In order to obtain the global characteristics of soil temperature, we connect the previous steps to the output layer directly, and the local characteristics of soil temperature are obtained through the following steps. We set different steps to the channels; with the outputs as cells, the estimation result is calculated by fully connecting different cells, using the past T_s data from the Laegern and Fluehli stations in Switzerland from 2006 to 2014 to estimate T_s in the next 6 hrs, 12 hrs, and 24 hrs at different soil depths (5, 10, and 15 cm).

The main contributions of this paper for T_s estimation are listed as follows:

- (1) According to our research, the GRU network had not been used for soil temperature prediction yet, and the method based on GRU was achieved in this paper for the purpose of estimating soil temperature.
- (2) In order to obtain the global characteristics of soil temperature, we connect the previous steps to the output layer directly, and the local characteristics of soil temperature are obtained through the following steps.
- (3) As the results showed, our method has a better performance than the other advanced technology available.

2. Materials and Methods

2.1. The Framework for Soil Temperature Estimation. First of all, we extract the corresponding past T_s data from the Laegern and Fluehli stations on FLUXNET as the input to our model. Meanwhile, we consider several other models based on machine learning technology models (LSTM, BPNN, ELM, and GRU). In our model, we connect the previous steps to the output layer directly to predict T_s . Finally, we compare the results by several statistical evaluation criteria (RMSE, MAE, MSE, and R^2) to evaluate the performance of the model. Figure 1 shows the overall structure of our study.

2.2. The Structure of GRU. The GRU network has the characteristics of simple structure and fast training speed and can transmit relevant information to the time series for prediction. It is widely applied in many fields due to its advantages precisely. The GRU can solve the time series problem and the gradient problem in backpropagation. The GRU unit structure is shown in Figure 2. It has two gates, which are the reset gate and the update gate. The update gate decides which new information should be discarded and added. The reset gate is used to decide how much past information to forget.

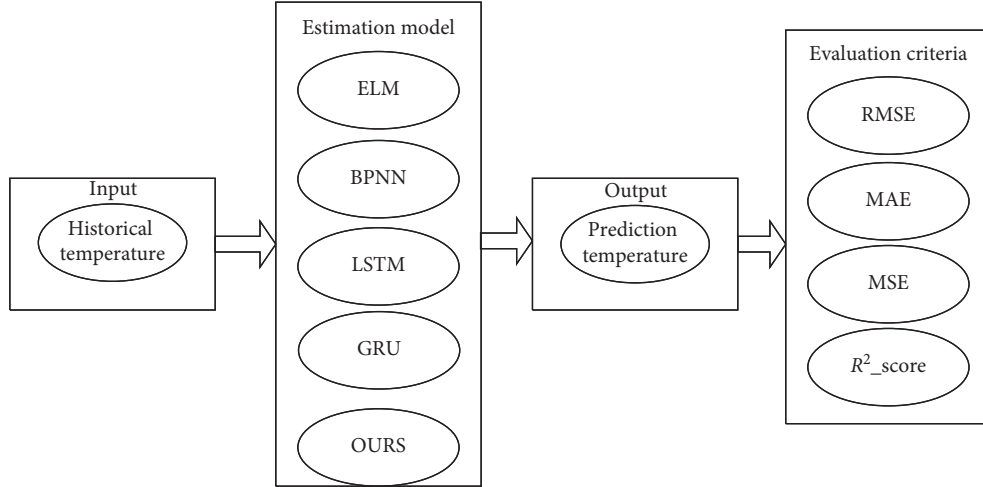


FIGURE 1: The overall structure of our study.

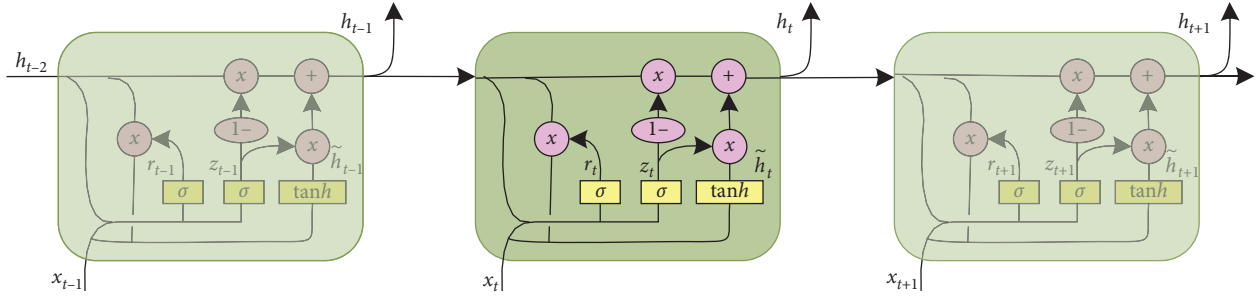


FIGURE 2: The structure of an GRU.

The calculation formulas of the GRU are as follows:

$$\begin{aligned}
 r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]), \\
 z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]), \\
 \tilde{h}_t &= \tanh\left(W_{\tilde{h}} \cdot [r_t * h_{t-1}, x_t]\right), \\
 h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t, \\
 y_t &= \sigma(W_0 \cdot h_t),
 \end{aligned} \quad (1)$$

where $x(t)$ represents the input value at the current moment and h_{t-1} is the hidden state of the previous node and uses them to get the gate status. r_t is the reset gate, z_t is the reset gate, \tilde{h}_t represents the new memory, h_t represents the hidden state, y_t is the output of the output layer, $\sigma(\cdot)$ is the sigmoid activation function, and $\tanh(\cdot)$ is the output tangent function.

2.3. The Structure of Our Model. Through the previous analysis, when there are a large number of cells in the GRU network, the correlation between the features will decrease with the time series extending. Therefore, we proposed the model based on GRU is to solve the problem and improve the accuracy of the estimation model, as the topological structure is shown in Figure 3.

Our model network is composed of the traditional GRU network and the auxiliary networks. The information is updated to the next cell through the parameter back-propagation of the hidden state in the GRU model that would decrease the correlation with the earlier. The traditional GRU network is used as the basic network to obtain local features, and the auxiliary network is composed of the output of different steps to obtain global features, meanwhile, merging the features as the output of the entire network model. Input the past T_s data into our model to learn the pattern of periodic changes in T_s , which can enhance the correlation between past T_s data and improve the accuracy for the prediction.

We proposed the final output of our model is $y_{\text{our method}}(t)$ at the time step t , which combined all the channels to the fully connected layer, as follows:

$$y_{\text{our method}}(t) = \sum_{ni=0}^n W_{ni} y(t - ni + 1)_{ni} + b. \quad (2)$$

2.4. Objective Function Optimization Algorithm. Adaptive moment estimation algorithm has excellent performance with high computational efficiency and low memory requirements [31]. It can calculate different adaptive learning

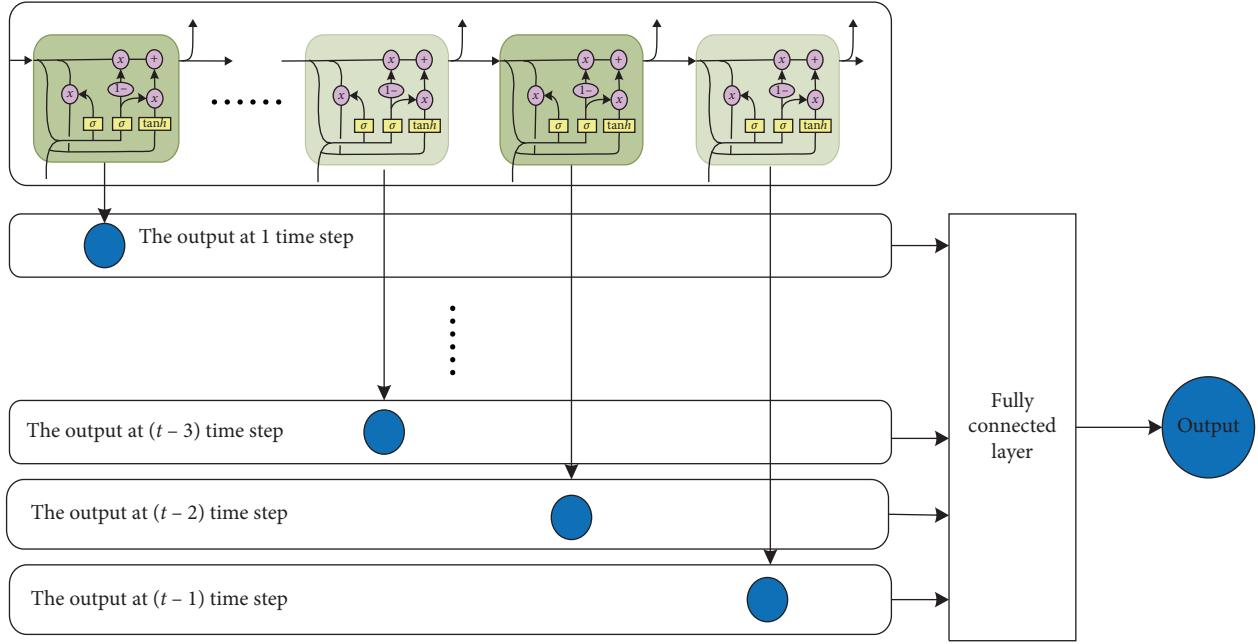


FIGURE 3: The structure of our model.

rates for different parameters. The method is suitable for processing large-scale data and optimizing parameters, as well as solving sparse gradient problems. The Adam is widely used in the field of deep learning, and it is used to optimize the model in this paper. This article uses the mean square error to optimize our model, and the calculation formula is as follows:

$$L = \frac{1}{N} \sum_{i=1}^N (y(t) - \hat{y}_{\text{our method}}(t))^2, \quad (3)$$

where the observed data from the stations is $y(t)$ and the output of our model is $\hat{y}_{\text{our method}}(t)$.

Learning rate is an important parameter in deep learning. The value of the learning rate will affect the convergence of the function. When the learning rate is set too small, the convergence will be very slow. Meanwhile, when the value is set too large, the gradient will be affected. We use the method of exponential decay learning rate to improve the convergence of Adam and the method of exponential decay learning rate. By constantly adjusting the learning rate, the step size is set to 100 and the attenuation rate is 0.96; the algorithm is close to the optimal solution.

2.5. Model Training and Test. In this paper, the past T_s data (the Laegern and Fluehli stations in Switzerland from 2006 to 2014) is served as the input to our model for estimating T_s and using TensorFlow backend. The model is tested on Intel Core (TM) i7-5820K, 3.30 GHz CPU, and 64 GB memory running Pycharm 2018. We use three-quarters of all data as training samples (data during 2006.1.1–2013.3.14), and the others were used as testing samples (data during 2013.3.15–2014.12.31). We assume that the value at the t point in the time series is x_t , which is predicted by the first

$t-1$ elements; use half of the daily soil temperature to predict the values of T_s in the following 6 hrs, 12 hrs, and 24 hrs; and set the value of t to 24. This paper compared our model with the other models (including ANN, LSTM, ELM, and GRU), meanwhile, calculating several evaluation criteria (RMSE, MAE, MSE, and R^2) to estimate the model performance, as follows:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{n=1}^N (y_i - \hat{y}_i)^2}{N}}, \\ \text{MAE} &= \frac{\sum_{n=1}^N |y_i - \hat{y}_i|}{N}, \\ \text{MSE} &= \frac{\sum_{n=1}^N (y_i - \hat{y}_i)^2}{N}, \\ R^2 &= \frac{\sum_{n=1}^N (y_i - \bar{y})^2 - \sum_{n=1}^N (y_i - \hat{y}_i)^2}{\sum_{n=1}^N (y_i - \bar{y})^2}, \end{aligned} \quad (4)$$

where the total number of data is denoted as N , y_i is the observed value of data at the moment, \hat{y}_i is the predicted value obtained through different methods, \bar{y} is the observed average value of the data. According to our knowledge, we can understand the fitting degree of the model and the accuracy of data prediction through evaluation criteria. With the smaller value of RSME, MAE, and MSE and the larger value of R^2 , the model will show the best performance.

2.6. Study Area and Field Experiment. This paper studied the data from two stations (Laegern 47.48 N, 8.37 E, Fluehli 46.88 N, 8.01 E) located in Switzerland and downloaded the past T_s data within half an hour on FLUXNET (<https://fluxnet.fluxdata.org/>) to verify our model. Since these two

stations are located in their domestic ecological nature reserve, T_s has a certain impact on the surrounding ecological environment, such as plant growth, soil fertility, and microbial activities, as shown in Figure 4.

This article takes the past T_s data from the stations as the input. With the data provided by the stations, it can be seen that the depth at the 15 cm soil temperature of the Laegern station is the most stable, achieving the minimum of temperature differences and standard deviation; x_{\min} is the minimum value; x_{\max} is the maximum value; x_{mean} is the average value; z_{sd} represents the standard deviation; z_s represents skewness; and z_v represents variation coefficient which is shown in Table 1.

3. Results and Discussions

In this paper, comparing our model with the other four models (BPNN, LSTM, ELM, GRU) for estimating T_s at the two stations' data, use the Adam to optimize the model and experiment with scikit-learn. For instance, input layer, hidden layer, and output layer consist of the ANN. We set the batch size to 10000, the number of iterations to 100, the learn rate to 0.03, and the number of nodes to 32, and the model get the best performance. We apply the elm function to the ELM model, the activation function in the hidden layer is sigmoid, and the number of nodes is set same to ANN. We Set the same hyperparameters to the GRU and our model so that it can be useful to show the performance of our model.

3.1. Evaluation for the Hyperparameters in Our Model. According to our research, the value of the hyperparameters has a certain influence on the performance of the model. As the hyperparameters, what we mentioned are the number of channels, iterations, learning rate, and the number of nodes ($\text{num}_{\text{our model}}$). For example, we estimate the soil temperature at the depth of 5 cm for 6 hrs of the Laegern station. As the results showed, $\text{num}_{\text{our model}}$ has a certain impact on the fitting of the model and the acquisition of relevant important information. With the learning rate set to 0.03, the number of nodes is 32, the number of channels is set to 4, the number of iterations is set to 100, and our model has the best performance.

When the model is overfitting during the training process, it is not conducive to the model adapting to the changes of the data. In contrast, if underfitting occurs during training, it is not conducive to data mining of related data. Moreover, due to the large learning rate value for optimal weights, the predictive model can easily be trapped into local optimum during the learning process. Then, when the learning rate value is small, it will make the parameters hardly converge to the optimal value for training the predictive model. The results are shown in Table 2, as follows:

3.2. Evaluation for Different Models. Comparing our model with the other four models (BPNN, ELM, LSTM, and GRU) in this part. The inputs to all the predictive models were the past T_s data from the stations. The output of the

model was estimated T_s values in the following 6 hrs, 12 hrs, and 24 hrs.

The predicting results of the five models at the depths of 5, 10, and 15 cm in the following 6 hrs, 12 hrs, and 24 hrs are experimented for the Laegern station as shown in Table 3. For the depth of 5 cm results in the coming 6 hrs, our model has better performance than the other models. According to the results, our model is compared with other models, and the percentage of the RMSE has reductions of 35.4% (GRU), 38.9% (LSTM), 39.1% (ELM), and 53.4% (BPNN), respectively; the percentage of the MAE has reductions of 42.9% (GRU), 46.2% (LSTM), 45.4% (ELM), and 58.2% (BPNN); and the percentage of the MSE has reductions of 57.3% (GRU), 58.6% (LSTM), 58.8% (ELM), and 75.9% (BPNN), getting the R^2 score amount to 0.9638 (our model) compared to 0.9034 (LSTM), 0.9022 (ELM), 0.8319 (BPNN), and 0.9058 (GRU). Because our model has the lower value of RMSE, MAE, and MSE, meanwhile, having the higher value of R^2 , it is obvious that our model has a very excellent performance on T_s estimation. However, for the depth of 5 cm results in the following 12 hrs and 24 hrs, we can draw the same conclusion that our model performance is still stronger than the other four models. The accuracy has been continuously improved from 5 cm to 15 cm of the soil depths, but the accuracy in the following 6 hrs to 24 hrs has been decreased. It may be caused by systematic errors when the model makes long-term predictions [32]. However, the ELM model gets a superior level of accuracy compared to the others models for the depth of 5 cm in the following 12 hrs, the depth of 10 cm in the following 6 hrs, and the depth of 15 cm. The reason for this might be ELM is a feedforward neural network architecture in which parameters are randomly chosen [19]. Sometimes a nonoptimal solution may be generated, which will affect the performance of the model.

We estimate the soil temperature at the depth of 5 cm, 10 cm, and 15 cm of the Laegern station in the following 6 hrs, 12 hrs, and 24 hrs, respectively. The estimation results are at the depth of 5 cm, 10 cm, and 15 cm of the Laegern station in the following 6 hrs, 12 hrs, and 24 hrs. The linear relationship between the estimated value and the observed value is shown in Figure 5. According to the distribution of the scatter plot, it can be seen that the linear relationship of our model is closer to the ideal line ($y=x$), and its R^2 value is higher than the others. It is shown in Figures 5(a) and 5(b), for instance, that the higher value of R^2 in our model is 0.9638 for the depth of 5 cm in the following 6 hrs, and the linear relationship is $y=1.0063x+0.0438$. However, all the tested models get a good performance when we estimate at the depth of 15 cm in the following 6 hrs and 12 hrs, which is shown in Figure 5(c). That is because the estimated value and the observed value have better consistency in this case. As the time series extends (24 hrs), our model still maintains good accuracy, and the accuracy of other models starts to decrease. Above all, the experimental results showed that our model has a certain degree of robustness for long-term estimation.

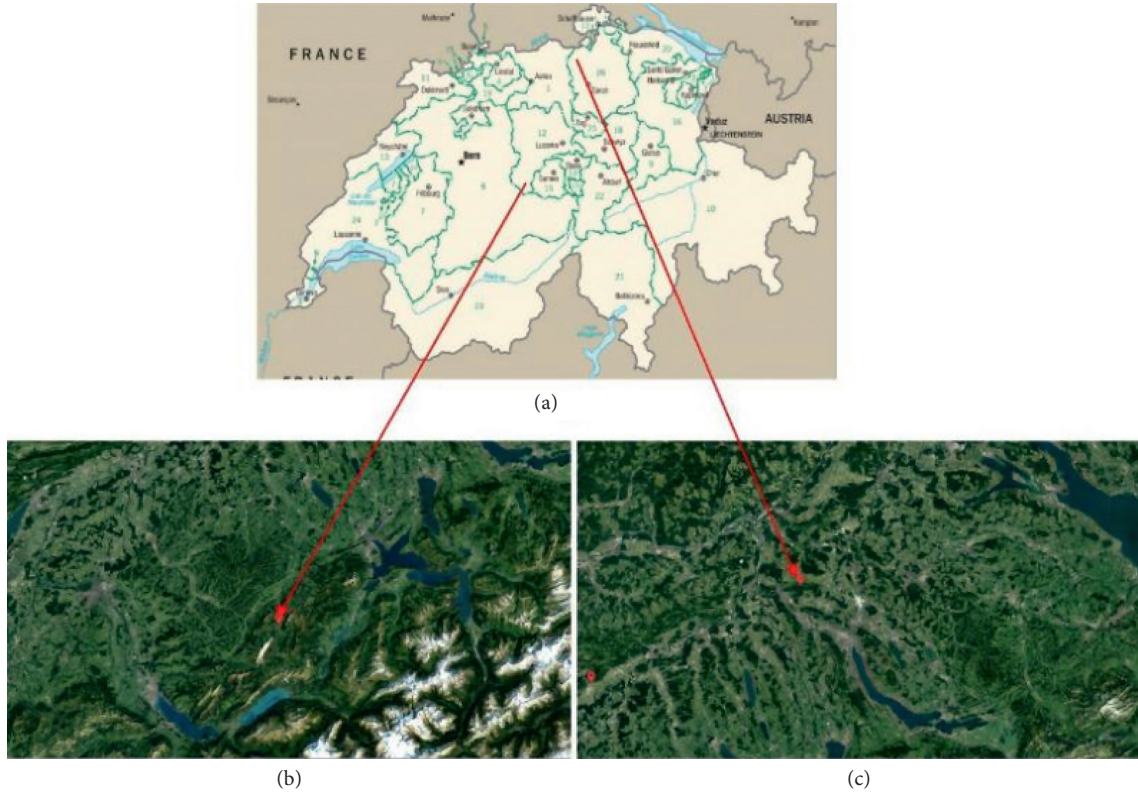


FIGURE 4: Location and geographical information of the experimental station. (a) The experimental station located in Switzerland. (b) Satellite image of the Fluehli station at 10 km scale. (c) Satellite image of the Laegern station at 10 km scale. The satellite images were obtained from google earth.

TABLE 1: Statistical properties of the meteorological information and soil temperature data for Laegern and Fluehli stations.

Station	Variable($^{\circ}\text{C}$)	x_{\min}	x_{\max}	x_{mean}	z_{sd}	z_s	z_v
Laegern	T_s -5 cm	-1.888	26.876	10.104	6.061	0.103	0.599
	T_s -10 cm	-0.181	22.193	9.726	5.435	-0.031	0.558
	T_s -15 cm	0.16	19.394	9.010	5.025	-0.068	0.557
Fluehli	T_s -5 cm	-0.35	21.822	8.729	6.338	0.075	0.726
	T_s -10 cm	-0.044	21.727	8.836	6.242	0.071	0.706
	T_s -15 cm	0.432	20.826	8.813	6.023	0.062	0.683

TABLE 2: Results of Laeger's T_s , with different hyperparameters set to our model.

Learning rate	$\text{num}_{\text{our model}}$	n	Iterations	R^2	Estimation time (second)
0.03	32	4	100	0.9691	48.790
0.03	32	4	200	0.9692	99.738
0.03	32	4	50	0.9612	22.534
0.03	64	4	100	0.9676	91.368
0.03	16	4	100	0.9623	30.257
0.003	32	4	100	0.9615	49.391
0.3	32	4	100	0.9668	48.006
0.03	32	3	100	0.9683	41.339
0.03	32	5	100	0.9679	55.641

The frequency plot of the absolute estimation error is shown in Figure 6, and each bar indicates the error percentage. It can be obviously seen that our model has the highest frequency for the smallest error magnitude encountered in

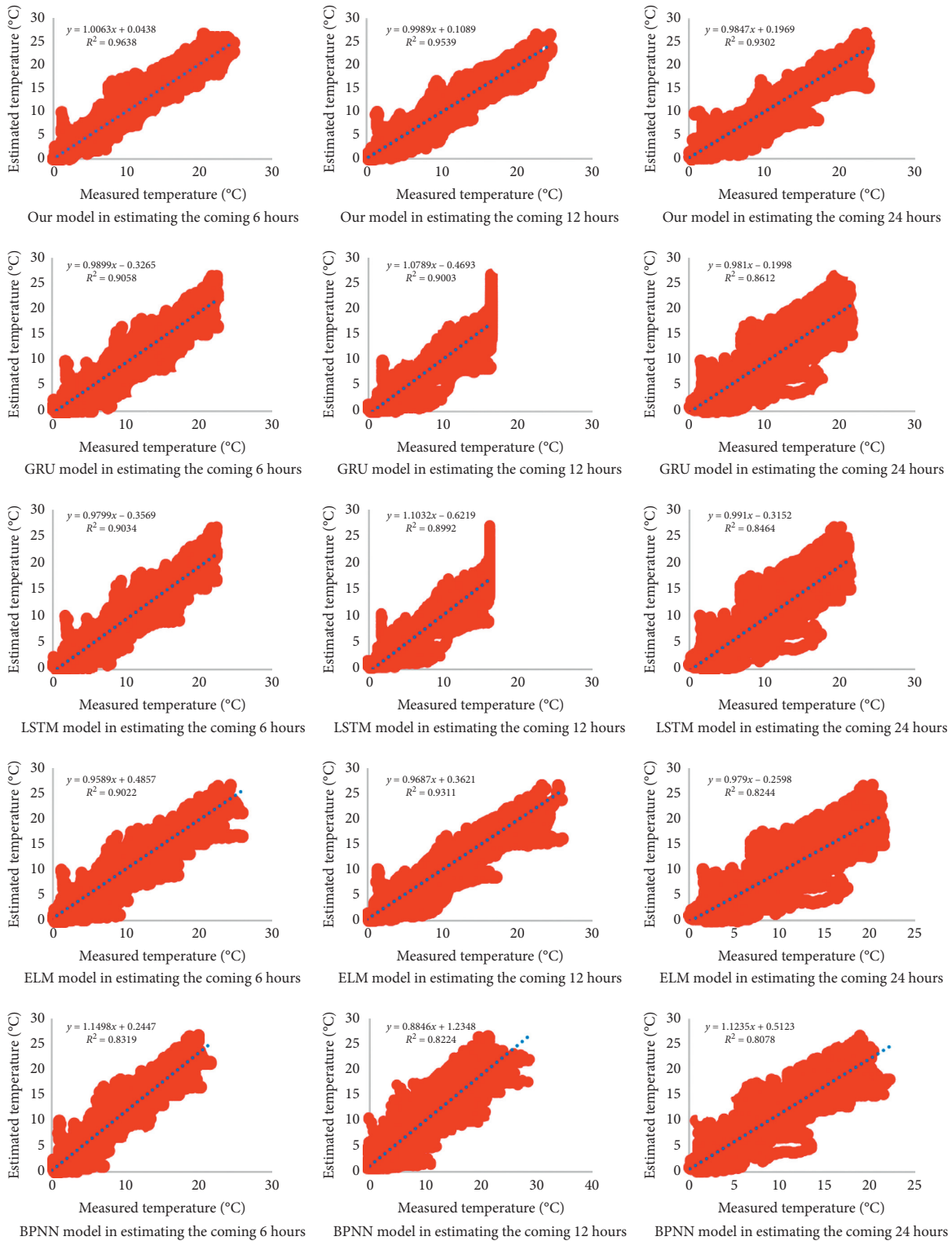
estimating T_s of the Laegern station. For example, for the depth of 5 cm in the following 6 hrs, our model has a higher value of frequency (73.8%) compared to the other four models (46.5% (GRU), 45.9% (LSTM), 46.6% (ELM), 36.8% (BPNN)).

TABLE 3: The testing phase results of the different models at Laegern station. Note that the optimal models are boldfaced.

Depth (cm)	Estimation (hrs)	Method	RMSE	MAE	MSE	R^2	
5	6	BPNN	2.3634	1.8168	5.5887	0.8319	
		ELM	1.8075	1.3898	3.2672	0.9022	
		LSTM	1.8037	1.4109	3.2542	0.9034	
		GRU	1.7054	1.3285	3.1537	0.9058	
			Our model	1.1012	0.7586	1.3458	0.9638
	12	BPNN	2.4358	1.8543	5.9368	0.8224	
		ELM	1.5141	1.1219	2.2926	0.9311	
		LSTM	1.8387	1.3191	3.3848	0.8992	
		GRU	1.7321	1.2229	3.2789	0.9003	
			Our model	1.2576	0.9954	1.7438	0.9539
	24	BPNN	2.5215	1.8938	6.3525	0.8078	
		ELM	2.4164	1.7991	5.8390	0.8244	
LSTM		2.2669	1.7470	5.1438	0.8464		
GRU		2.1569	1.6421	5.0349	0.8612		
		Our model	1.4221	1.0287	2.2316	0.9302	
10	6	BPNN	1.3643	1.1015	1.8660	0.9308	
		ELM	0.8011	0.6077	0.6418	0.9764	
		LSTM	0.8015	0.6268	0.6482	0.9743	
		GRU	0.7065	0.5343	0.5489	0.9782	
			Our model	0.4238	0.2756	0.1714	0.9905
	12	BPNN	1.2214	0.8875	1.5035	0.9476	
		ELM	0.8296	0.6273	0.6883	0.9747	
		LSTM	0.8085	0.6134	0.6551	0.9773	
		GRU	0.7998	0.5946	0.5589	0.9765	
			Our model	0.6203	0.4421	0.4212	0.9856
	24	BPNN	1.3445	1.0136	1.8175	0.9346	
		ELM	1.2105	0.9174	1.4652	0.9460	
LSTM		1.1412	0.8756	1.2989	0.9518		
GRU		1.0462	0.7216	1.1236	0.9563		
		Our model	0.7658	0.5554	0.6513	0.9756	
15	6	BPNN	0.9325	0.8367	0.8734	0.9659	
		ELM	0.5790	0.4275	0.3353	0.9858	
		LSTM	0.5747	0.4589	0.3363	0.9851	
		GRU	0.4825	0.3546	0.2445	0.9868	
			Our model	0.2754	0.1643	0.1287	0.9951
	12	BPNN	0.8812	0.7145	0.7785	0.9654	
		ELM	0.6626	0.5008	0.4390	0.9814	
		LSTM	0.6564	0.5175	0.4225	0.9868	
		GRU	0.5638	0.4089	0.3896	0.9865	
			Our model	0.4755	0.3241	0.2114	0.9875
	24	BPNN	0.9175	0.6842	0.8318	0.9653	
		ELM	0.9024	0.6892	0.8143	0.9654	
LSTM		0.8984	0.7052	0.8068	0.9646		
GRU		0.7968	0.6068	0.6994	0.9689		
		Our model	0.6926	0.5435	0.4797	0.9799	

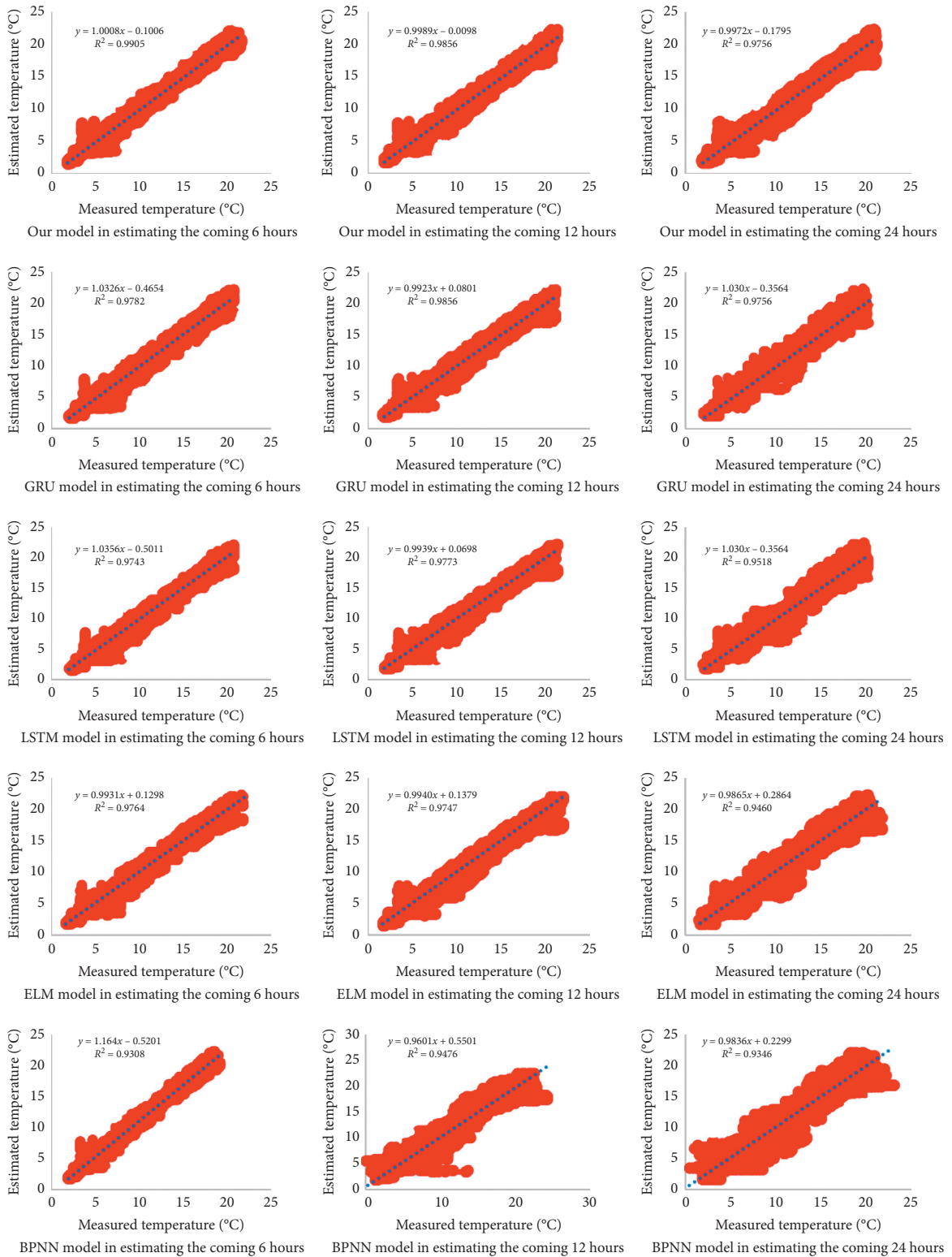
Using different predictive models to test the data separately from the Fluehli station at the depths of 5 cm, 10 cm, and 15 cm in the following 6 hrs, 12 hrs, and 24 hrs, the results are shown in Table 4. The results noted that our model mainly gets better performance than the others; for example, the proposed model achieved excellent results

(RMSE = 0.6534, MAE = 0.3928, MSE = 0.4735, and $R^2 = 0.9860$) compared to the other four models at the depth of 5 cm in the following 6 hrs at Fluehli station. Generally speaking, our model mainly has superior performance for estimating T_s with the experiments in different regions, different times, and different soil depths.



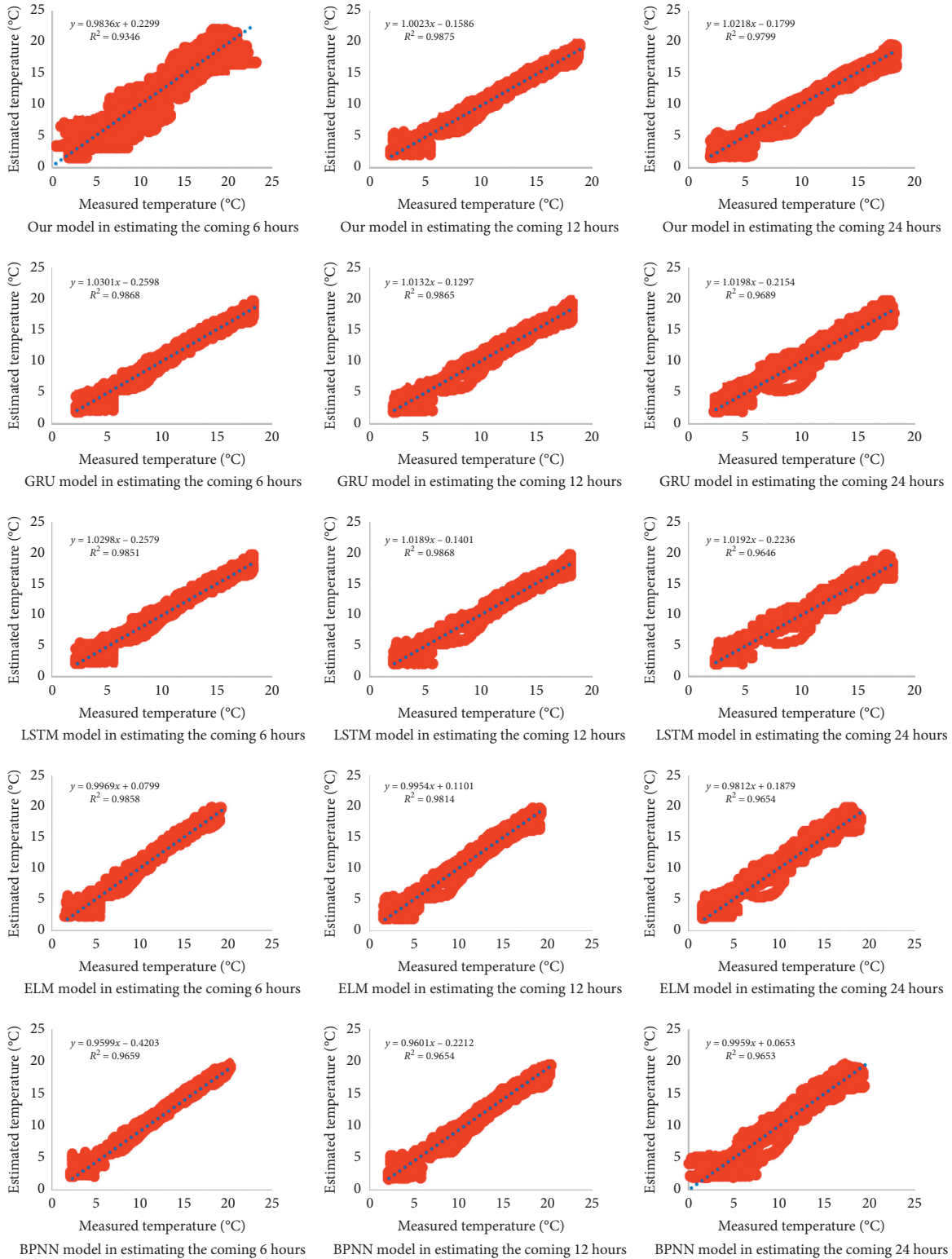
(a)

FIGURE 5: Continued.



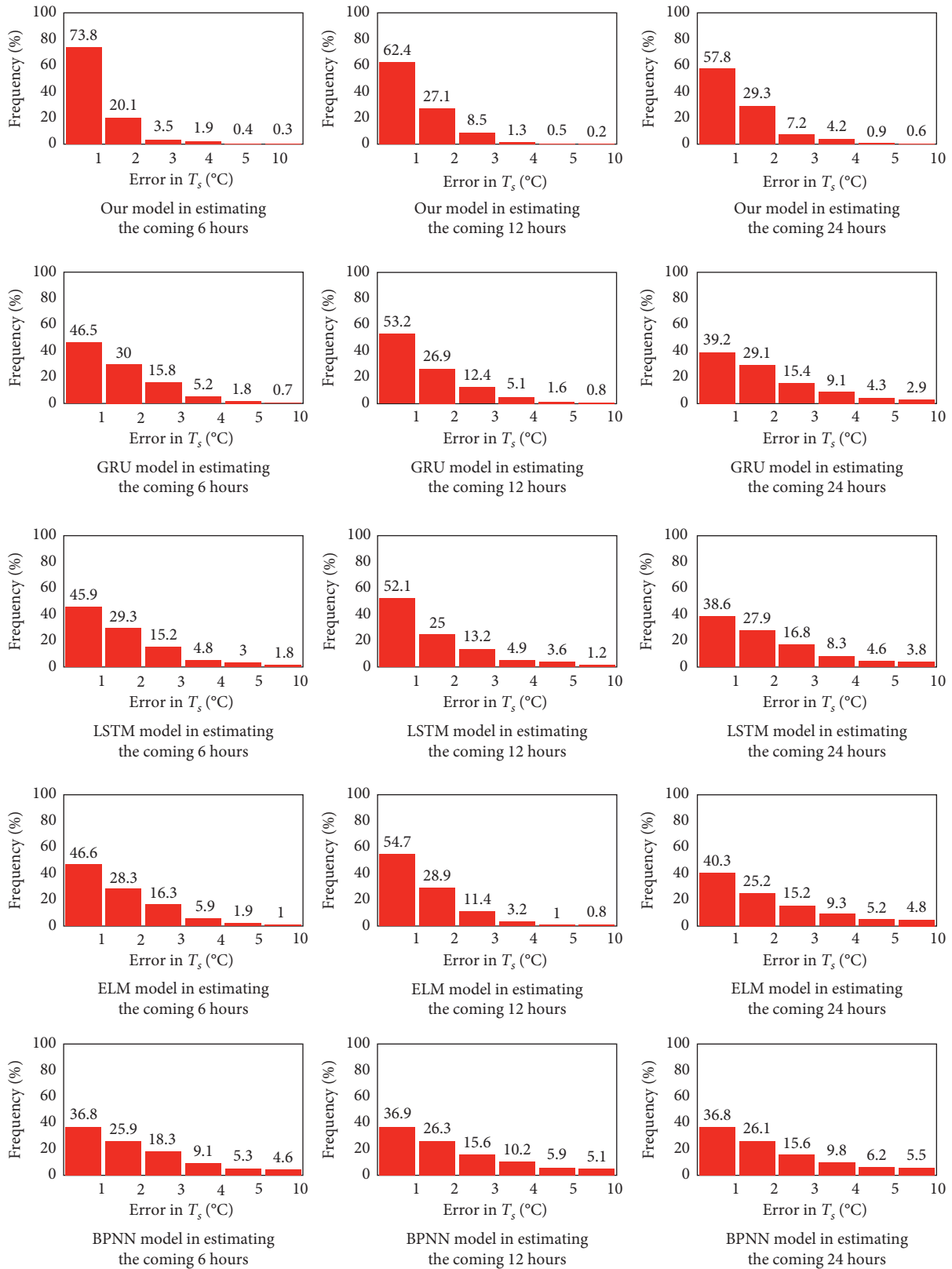
(b)

FIGURE 5: Continued.



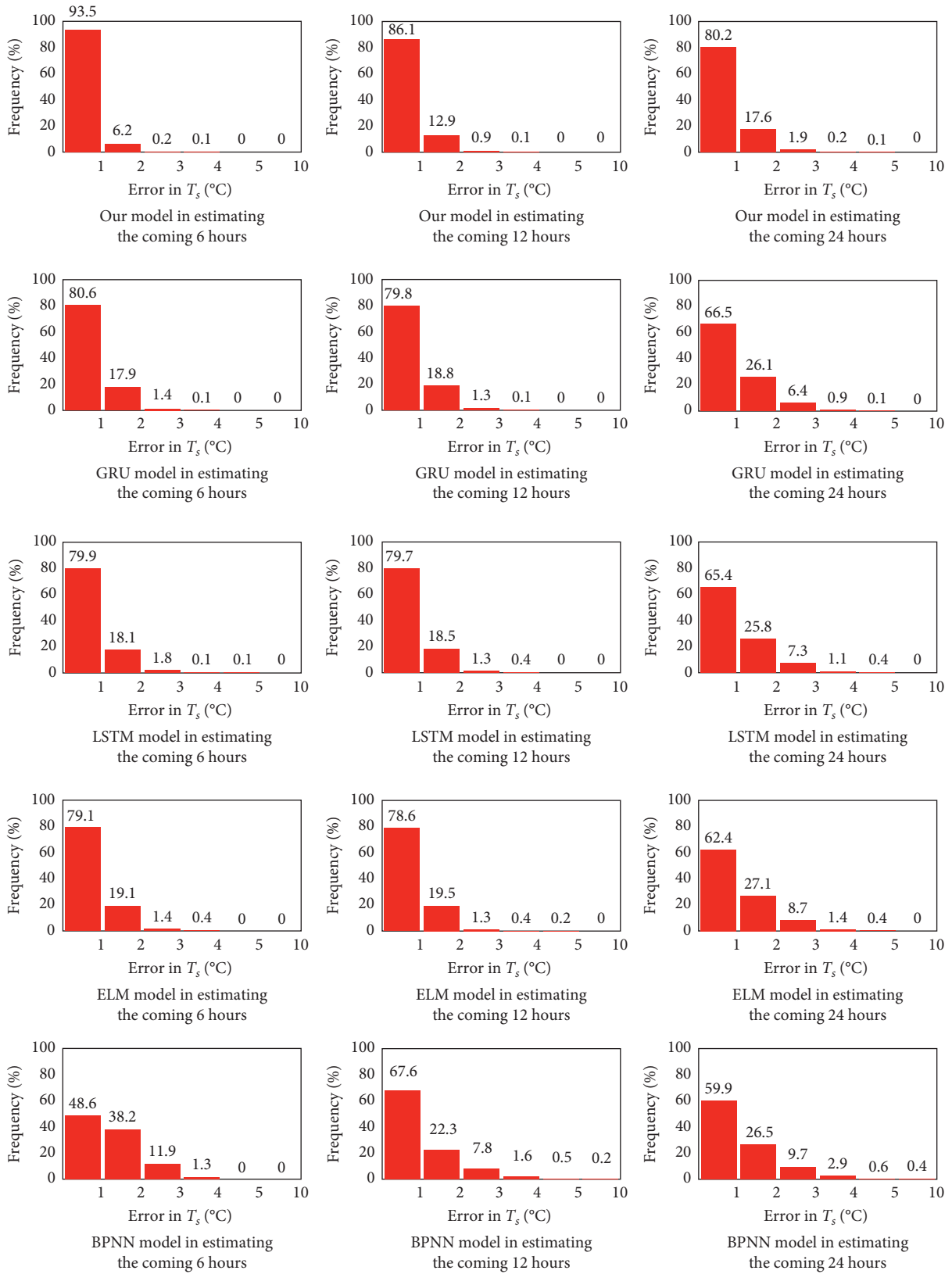
(c)

FIGURE 5: The scatterplots of estimated and observed soil temperature using different predictive models for the Laegern station. (a) The depth of 5 cm. (b) The depth of 10 cm. (c) The depth of 15 cm.



(a)

FIGURE 6: Continued.



(b)

FIGURE 6: Continued.

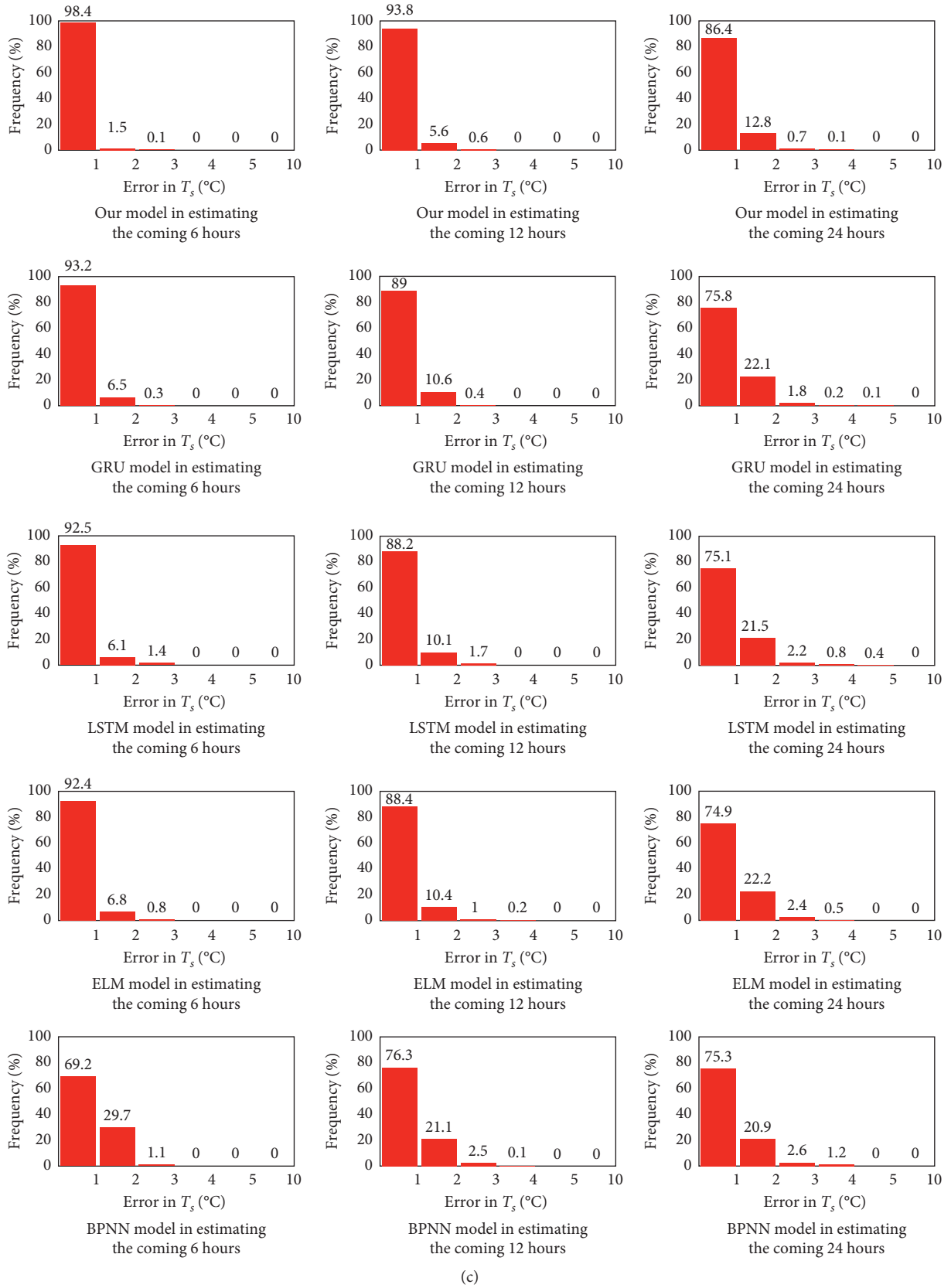


FIGURE 6: The frequency plot of the absolute estimation error using the predictive models for the Laegern station. (a) The depth of 5 cm. (b) The depth of 10 cm. (c) The depth of 15 cm.

TABLE 4: The testing phase results of the different models at Fluehli station. Note that the optimal models are boldfaced.

Depth (cm)	Estimation (hrs)	Method	RMSE	MAE	MSE	R^2		
5	6	BPNN	0.9364	0.6339	0.8779	0.9758		
		ELM	1.2822	0.9026	1.6441	0.9556		
		LSTM	0.8759	0.5731	0.7706	0.9787		
		GRU	0.7689	0.4125	0.6758	0.9799		
			Our model	0.6534	0.3928	0.4735	0.9860	
	12	6	BPNN	1.1428	0.8079	1.3054	0.9612	
			ELM	1.0470	0.7191	1.0962	0.9703	
			LSTM	1.0378	0.7186	1.0801	0.9712	
			GRU	0.9989	0.6838	0.9945	0.9756	
				Our model	0.9276	0.6149	0.9587	0.9789
		24	BPNN	1.7823	1.2545	3.1931	0.9158	
			ELM	1.7248	1.1940	2.9748	0.9169	
LSTM			1.1947	0.8367	1.4223	0.9646		
GRU	1.0845		0.7912	1.3213	0.9698			
		Our model	0.9753	0.6456	1.0627	0.9763		
10	6	BPNN	1.8578	1.2665	3.4347	0.9012		
		ELM	1.0416	0.7248	1.0849	0.9694		
		LSTM	1.0258	0.6774	1.0482	0.9713		
		GRU	0.9846	0.5678	0.9312	0.9756		
			Our model	0.4668	0.2585	0.2214	0.9914	
	12	6	BPNN	1.6376	1.1389	2.6865	0.9237	
			ELM	0.9267	0.6348	0.8587	0.9757	
			LSTM	0.9325	0.6513	0.8645	0.9752	
			GRU	0.8278	0.5498	0.7723	0.9787	
				Our model	0.8036	0.5203	0.7213	0.9796
		24	BPNN	1.1402	0.7924	1.3038	0.9626	
			ELM	1.6616	1.1507	2.7610	0.9218	
LSTM			1.1042	0.7657	1.2347	0.9642		
GRU	1.0036		0.6738	1.1289	0.8599			
		Our model	0.8423	0.5523	0.7589	0.9742		
15	6	BPNN	1.0816	0.8047	1.1658	0.9623		
		ELM	0.7741	0.5640	0.5992	0.9822		
		LSTM	0.7565	0.5259	0.5753	0.9876		
		GRU	0.6489	0.4376	0.4978	0.9889		
			Our model	0.4099	0.3254	0.2512	0.9910	
	12	6	BPNN	0.9703	0.7025	0.9449	0.9736	
			ELM	0.7085	0.5033	0.5020	0.9851	
			LSTM	0.7123	0.5046	0.5012	0.9846	
			GRU	0.6368	0.4532	0.4465	0.9856	
				Our model	0.5656	0.3984	0.3979	0.9878
		24	BPNN	1.3065	0.9434	1.6957	0.9461	
			ELM	1.1466	0.8188	1.3147	0.9609	
LSTM			0.8836	0.6249	0.7779	0.9746		
GRU	0.7946		0.5639	0.6842	0.9755			
		Our model	0.6239	0.4183	0.4203	0.9868		

4. Conclusions

Soil temperature (T_s) as the important variable is one of the land surface features impact on Earth science, usually used in many research fields; for example, it affects the growth and development of plants and the formation of soil. In this study, we research the performance of backpropagation neural networks (BPNN), gated recurrent unit (GRU), extreme learning machine (ELM), long short-term memory (LSTM) network, and our model for estimating T_s at the depth of 5 cm, 10 cm, and 15 cm in the following 6 hrs, 12 hrs, and 24 hrs over the Switzerland Laegern and Fluehli

stations. The statistical results indicated that our model mainly performs better than the other four models on the T_s estimation.

In order to reduce the influence of long-term series on the accuracy of soil temperature estimation and obtain the global characteristics of soil temperature, we connect the previous steps to the output layer directly, and the local characteristics of soil temperature are obtained through the following steps.

The soil temperature is affected by many factors, such as atmospheric temperature and precipitation. EEMD can decompose time series into signals of different frequencies

and display the time series values with distinguishing ability under different frequencies. This method may decompose the complex soil temperature time series into discriminative data. Therefore, we will attempt to integrate EEMD to decompose the soil temperature time series and the features with different frequencies into the model to improve estimation accuracy.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant no. 51805203, the Science and Technology Development Plan of Jilin Province 20190201023JC, and the Development and Reform Commission of Jilin Province (2019C054-2).

References

- [1] L. Lai, X. Zhao, L. Jiang et al., "Soil respiration in different agricultural and natural ecosystems in an arid region," *PLoS One*, vol. 7, no. 10, pp. e48011–9, 2012.
- [2] H. Tabari, P. Hosseinzadeh Talaei, and P. Willems, "Short-term forecasting of soil temperature using artificial neural network," *Meteorological Applications*, vol. 22, no. 3, pp. 576–585, 2014.
- [3] A. Keshavarzi, F. Sarmadian, E.-S. E. Omran, and M. Iqbal, "A neural network model for estimating soil phosphorus using terrain analysis," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 18, no. 2, pp. 127–135, 2015.
- [4] O. Kisi and H. Sanikhani, "Modelling long-term monthly temperatures by several data-driven methods using geographical inputs," *International Journal of Climatology*, vol. 35, no. 13, pp. 3834–3846, 2015.
- [5] M. Saeid, B. Javad, and K. Keivan, "Evaluating the performance of artificial intelligence methods for estimation of monthly mean soil temperature without using meteorological data," *Environmental Earth Sciences*, vol. 76, no. 8, p. 325, 2017.
- [6] L. L. Liang, D. A. Riveros-Iregui, R. E. Emanuel, and B. L. McGlynn, "A simple framework to estimate distributed soil temperature from discrete air temperature measurements in data-scarce regions," *Journal of Geophysical Research: Atmosphere*, vol. 119, no. 2, pp. 407–417, 2014.
- [7] S. Mehdizadeh, J. Behmanesh, and K. Khalili, "Evaluating the performance of artificial intelligence methods for estimation of monthly mean soil temperature without using meteorological data," *Environmental Earth Sciences*, vol. 76, no. 8, p. 325, 2017.
- [8] J. Qi, S. Li, Q. Li, Z. Xing, C. P.-A. Bourque, and F.-R. Meng, "A new soil-temperature module for SWAT application in regions with seasonal snow cover," *Journal of Hydrology*, vol. 538, pp. 863–877, 2016.
- [9] Y. Liu, A. H. Weerts, M. Clark et al., "Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities," *Hydrology and Earth System Sciences*, vol. 16, no. 10, pp. 3863–3887, 2012.
- [10] J. Shiri, S. Kim, and O. Kisi, "Estimation of daily dew point temperature using genetic programming and neural networks approaches," *Hydrology Research*, vol. 45, no. 2, pp. 165–181, 2014.
- [11] O. Kisi and H. Sanikhani, "Prediction of long-term monthly precipitation using several soft computing methods without climatic data," *International Journal of Climatology*, vol. 35, no. 14, pp. 4139–4150, 2015.
- [12] K. Mohammadi, S. Shamsirband, A. Kamsin, P. C. Lai, and Z. Mansor, "Identifying the most significant input parameters for predicting global solar radiation using an ANFIS selection procedure," *Renewable and Sustainable Energy Reviews*, vol. 63, pp. 423–434, 2016.
- [13] M. A. Ghorbani, S. Shamsirband, D. Zare Haghi, A. Azani, H. Bonakdari, and I. Ebtehaj, "Application of firefly algorithm-based support vector machines for prediction of field capacity and permanent wilting point," *Soil and Tillage Research*, vol. 172, pp. 32–38, 2017.
- [14] M. Delbari, S. Sharifazari, and E. Mohammadi, "Modeling daily soil temperature over diverse climate conditions in Iran—a comparison of multiple linear regression and support vector regression techniques," *Theoretical and Applied Climatology*, vol. 135, no. 3–4, pp. 991–1001, 2019.
- [15] Y. Feng, N. Cui, W. Hao, L. Gao, and D. Gong, "Estimation of soil temperature from meteorological data using different machine learning models," *Geoderma*, vol. 338, no. 2019, pp. 67–77, 2019.
- [16] Q. Li, H. Hao, Y. Zhao et al., "GANs-LSTM model for soil temperature estimation from meteorological: a new approach," *IEEE Access*, vol. 8, pp. 59427–59443, 2020.
- [17] O. Kisi, M. Tombul, and M. Z. Kermani, "Modeling soil temperatures at different depths by using three different neural computing techniques," *Theoretical and Applied Climatology*, vol. 121, no. 2015, pp. 377–387, 2015.
- [18] H. Tabari, A.-A. Sabziparvar, and M. Ahmadi, "Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region," *Meteorology and Atmospheric Physics*, vol. 110, no. 3–4, pp. 135–142, 2011.
- [19] H. Sanikhani, R. C. Deo, Z. M. Yaseen, O. Eray, and O. Kisi, "Non-tuned data intelligent model for soil temperature estimation: a new approach," *Geoderma*, vol. 330, no. 2018, pp. 52–64, 2018.
- [20] S. Samadianfard, E. Asadi, S. Jarhan et al., "Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths," *Soil and Tillage Research*, vol. 175, no. 2018, pp. 37–50, 2018.
- [21] O. Kisi, H. Sanikhani, and M. Cobaner, "Soil temperature modeling at different depths using neuro-fuzzy, neural network, and genetic programming techniques," *Theoretical and Applied Climatology*, vol. 129, no. 3–4, pp. 833–848, 2017.
- [22] M. Zeynoddin, H. Bonakdari, I. Ebtehaj, F. Esmailbeiki, B. Gharabaghi, and D. Zare Haghi, "A reliable linear stochastic daily soil temperature forecast model," *Soil and Tillage Research*, vol. 189, no. 2019, pp. 73–87, 2019.
- [23] M. Bilgili, "The use of artificial neural networks for forecasting the monthly mean soil temperatures in Adana, Turkey," *Turkish Journal of Agriculture and Forestry*, vol. 35, no. 2011, pp. 83–93, 2011.

- [24] S. Mehdizadeh, F. Fathian, M. J. S. Safari, and A. Khosravid, "Developing novel hybrid models for estimation of daily soil temperature at various depths," *Soil and Tillage Research*, vol. 197, no. 2020, Article ID 104513, 2020.
- [25] A. Danandeh Mehr, E. Kahya, and E. Kahya, "A Pareto-optimal moving average multigene genetic programming model for daily streamflow prediction," *Journal of Hydrology*, vol. 549, pp. 603–615, 2017.
- [26] S. Bahrami, F. Doulati Ardejani, and E. Baafi, "Application of artificial neural network coupled with genetic algorithm and simulated annealing to solve groundwater inflow problem to an advancing open pit mine," *Journal of Hydrology*, vol. 536, pp. 471–484, 2016.
- [27] K. Cho, B. V. Merriënboer, C. Gulcehre et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, <https://arxiv.org/abs/1406.1078>.
- [28] H. Liu, J. Zhou, Y. Zheng, Y. Jiang, and Y. Zhang, "Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders," *ISA Transactions*, vol. 77, pp. 167–178, Article ID S0019057818301514, 2018.
- [29] S. Miao and W.-H. Hung, "River flooding forecasting and anomaly detection based on deep learning," *IEEE Access*, vol. 8, pp. 198384–198402, 2020.
- [30] F. Rui, Z. Zuo, and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," in *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, IEEE, Wuhan, Hubei Province, China, November 2016.
- [31] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *Proceedings of the ICLR*, San Diego, CA, USA, May 2015.
- [32] K. Judd and M. Small, "Towards long-term prediction," *Physica D: Nonlinear Phenomena*, vol. 136, no. 1-2, pp. 31–44, 2000.

Research Article

Denoising Speech Based on Deep Learning and Wavelet Decomposition

Li Wang,¹ Weiguang Zheng ,² Xiaojun Ma,³ and Shiming Lin ^{4,5}

¹College of Chinese Literature and Media, Hubei University of Arts and Science, Xiangyang 441000, China

²School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin 541004, China

³Qinghai GLI Technology Limited, Xining 810001, China

⁴School of Informatics (National Demonstrative Software School), Xiamen University, Xiamen 361005, China

⁵Department of Computer Engineering, Changji University, Changji 831100, China

Correspondence should be addressed to Weiguang Zheng; weiguang.zheng@foxmail.com and Shiming Lin; xmuls@xmu.edu.cn

Received 7 May 2021; Revised 21 June 2021; Accepted 8 July 2021; Published 16 July 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Li Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The work proposed a denoising speech method using deep learning. The predictor and target network signals were the amplitude spectra of the wavelet-decomposition vectors of the noisy audio signal and clean audio signal, respectively. The output of the network was the amplitude spectrum of the denoised signal. Besides, the regression network used the input of the predictor to minimize the mean square error between its output and input targets. The denoised wavelet-decomposition vector was transformed back to the time domain by the output amplitude spectrum and the phase of the wavelet-decomposition vector. Then, the denoised speech was obtained by the inverse wavelet transform. This method overcame the problem that the frequency and time resolution of the short-time Fourier transform could not be adjusted. The noise reduction effect in each frequency band was improved due to the gradual reduction of the noise energy in the wavelet-decomposition process. The experimental results showed that the method has a good denoising effect in the whole frequency band.

1. Introduction

In the actual environment, speech signals are inevitably affected by the noises from the surrounding environment, transmission media, and electrical noise inside the communication equipment. These interferences greatly degrade the performance of the speech processing system and affect the quality of speech. Speech denoising aims to reproduce clean speech from noise-polluted signals, which is crucial for various applications, such as automatic speech recognition (ASR) and hearing aids. Several speech-denoising and speech-enhancement methods have been proposed based on the statistical difference between the speech and noise characteristics, including spectral subtraction [1], based estimation [2], Wiener filtering [3], subspace method [4], nonnegative matrix factorization (NMF) [5], and minimum mean square error (MMSE) [6].

Most of the filtering methods are limited to window-adding or masking operation in the frequency domain or

time domain due to the strong time-frequency coupling between speech signals and noises. It is difficult for these filtering methods to achieve effective signal-noise separation. As a nonlinear filter, the neural network was applied to this problem in the past, such as the early use of the shallow neural network (SNN) for speech-denoising study. However, the constraints on computing power and the size of training data lead to the implementations of relatively small neural networks, limiting denoising performance.

By learning a deep nonlinear network structure, deep learning has the following advantages: achieving the approximation of complex functions, representing the distributed representation of input data, and demonstrating its powerful ability to learn data and essential characteristics from a few sample sets. Meanwhile, it emphasizes the deep structure of the learning model. The current learning framework usually adopts a multilevel model. In this way, the training of the model relies on a large number of data sets, highlighting the importance of big data for a complete

and complex model. Deep learning also focuses on feature learning. Deep neural networks (DNNs) contain multiple nonlinear hiding layers, showing great potential to capture the complex relationship between noises and clean speeches. Many training algorithms have been proposed to train a deep network. DNNs have been applied to speech recognition [7], speech denoising [8], and speech separation [9].

Recently, Zhao et al. [10] used both convolutional and recurrent neural network architectures to exploit local structures in both the frequency and temporal domains for speech enhancement. Tan and Wang [11] combined the convolutional code-decoder (CED) and long short-term memory (LSTM) into the convolutional recurrent network (CRN) to achieve real-time monophonic speech enhancement. The proposed model is independent of noise and speaker. Moreover, the trainable parameters of CRN are much smaller. The full connection layer involved in deep neural networks (DNN) and convolutional neural networks (CNN) may not accurately describe the local information of the speech signal, especially for the high-frequency component. Therefore, Fu et al. [12] proposed an enhancement model of a full convolutional network (FCN) based on the original waveform. The system performs speech enhancement in an end-to-end manner, different from most existing denoising methods only dealing with amplitude spectrum.

Speech is a time-varying signal, in which usually changes occur at syllabic rates of 10 times/sec and exceeds the fixed time intervals of 10–30 msec. Short-time Fourier transform (STFT) is often used to analyze the speech on a time-frequency range [8, 9]. However, the window length of the STFT is fixed, that is, the time-domain resolution is fixed. According to the Heisenberg uncertainty principle, the frequency-domain resolution is also fixed. For a low-frequency signal, the time interval should be wider to determine the frequency better; however, for high-frequency signals, the time domain should be narrower to locate them better in the time domain. The resolution of STFT is not adjustable in the time domain and frequency domain, so it is not suitable for broadband analysis.

Wavelet analysis, developed in the 1980s, plays an important role in signal processing [13]. Wavelet transform (WT) has multiresolution and can adjust the window function adaptively according to the signal frequency. For low-frequency signals, WT provides low time-domain resolution and high-frequency domain resolution. For high-frequency signals, it provides high resolution in the time domain and low resolution in the frequency domain [14]. The wavelet transform coefficient reaches a maximum value in a certain region, and this point is called the modulus maximum of the wavelet transform in the region. The modulus maxima of useful signals in the multiresolution analysis increase with the decreased resolution; however, the modulus maxima of noisy signals in the multiresolution analysis decrease with the decreased resolution [15]. Threshold values are set according to the characteristics of useful signals and noise, and the wavelet coefficient is analyzed using this threshold value. When the wavelet coefficient is lower than this threshold value, the wavelet coefficient corresponds to a noise signal. In the wavelet

domain, the threshold is used to distinguish the useful signal from the noise signal. Finally, processed wavelet coefficients are reconstructed to obtain denoised signals [16].

The work proposed a speech denoising method based on deep learning. The predictor and target network signals were the amplitude spectrum of the wavelet-decomposition vector of the noisy audio signal and clean audio signal, respectively. The output of the network was the amplitude spectrum of the denoising signal. The output spectrum and the phase of the wavelet-decomposition vector were used to transform the denoised wavelet-decomposition vector back to the time domain. Then, the denoised speech was obtained by the inverse-wavelet transform. This method overcame the problem that the frequency and time resolution of STFT could not be adjusted.

2. General Theory

2.1. Short-Time Fourier Transform. STFT is widely used in speech analysis and processing, suitable for slow signal and time-varying signal spectrum analysis. In this method, the speech signal is first divided into frames, and then, the Fourier transform is carried out for each frame. Each frame of the speech signal can be intercepted from a variety of stationary signal waveforms, and the short-time spectrum of each frame of speech is an approximation of the spectrum value of the smooth signal waveform. Since the signal of each frame is short and stable, the Fourier transform of the frame signal is calculated to obtain the STFT:

$$\text{STFT}_x(t, f) = \int_{-\infty}^{\infty} x(t)h(t - \tau)e^{-j2\pi f\tau} d\tau, \quad (1)$$

where $\text{STFT}_x(t, f)$ is the coefficient of STFT. STFT is a function of time t and frequency f , which shows how the frequency of the speech signal changes with time.

According to the above STFT transformation, its inverse transformation can be defined as

$$x_t = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{STFT}_x(t', f')w(t - t')e^{-j2\pi f't'} dt' df', \quad (2)$$

where $w(t)$ is a window function. The longer window length means higher spectral resolution; however, the time resolution of the long window decreases correspondingly. Due to the contradiction between the time resolution and the frequency resolution, the practical operation should be based on the STFT analysis, and the appropriate window length should be determined.

2.2. Wavelet Transform. STFT is a windowed FT transform. FT is based on sinusoidal functions of different frequencies, so the signal is often decomposed into the superposition sum of sinusoidal waves of different frequencies. The wavelet transform replaces the infinitesimal trigonometric basis function with the wavelet basis of finite length and attenuation, thus locating frequency and time. The continuous wavelet transform (CWT) is the inner product of wavelet function $\phi(t)$ and square-integrable function $x(t)$ with good local properties in the time-frequency domain:

$$\begin{aligned} \text{CWT}_x(a, b) &= f, \\ \phi_{a,b} &= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \phi^* \left(\frac{t-b}{a} \right) dt, \end{aligned} \quad (3)$$

where $a > 0$ is the scale factor and b the displacement factor. The scale factor plays an important role in wavelet transform. When it is very small, it will show the details of the signal changing rapidly. When it is large, the wavelet is extended to show the coarse features of the signal. When $\phi(t)$ meets the admissibility condition, the inverse continuous wavelet transform (ICWT) is

$$x(t) = \frac{1}{C_\phi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{CWT}_x(a, b) \frac{1}{\sqrt{a}} \tilde{\phi} \left(\frac{t-b}{a} \right) \frac{1}{a^2} dt da, \quad (4)$$

where $\tilde{\phi}(t)$ is a dual function of $\phi(t)$ and C_ϕ an admissible constant. The data from CWT has large redundancy, which may not be suitable for DNN training for denoising speech. Discrete WT (DWT) uses filter banks to implement the Mallat algorithm. Figure 1 shows the three-level DWT, where cA1, cA2, and cA3 are approximate coefficients containing low-frequency information of the signal. cd1, cd2, and cd3 are detail coefficients and contain high-frequency information of the signal. c is the wavelet-decomposition vector; l is the bookkeeping vector containing the number of coefficients of each level.

2.3. Convolution Neural Networks for Deep Learning. A convolution neural network of deep learning is a deep-learning network generated on the theoretical basis of a neural network. The neural network is a fully connected network, that is, each neuron in the upper layer is connected to a neuron in the next layer. In this case, for multidimensional input information such as sound or image, the amount of information contained is relatively large; for the hidden layer, the traditional BP algorithm requires more weight parameters. The resulting slow training speed leads to more samples required for training. Overfitting is more likely to occur with insufficient training. In this way, the parameters learned are not universal, so they cannot represent and restore the input signal.

Ordinary neural network structure does not consider the characteristics of the input data. Even for a little change in the original data, the neural network does not take into account the data characteristics for optimized training. The neural network is fully connected, and all input data need to be considered; thus, it is impossible to identify and train the local regional features in the data.

Given the problems existing in the above ordinary neural network structure, the convolutional neural network transforms the ordinary neural network through local connection to feel the field of vision, weight sharing, and subsampling process through a local connection. It is used to learn features. Figure 2 shows the convolutional neural network model.

The total core operation of convolution in the convolution layer is as follows:

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l \right), \quad (5)$$

where k is the convolution kernel (filter), l is the number of layers, M is the j^{th} feature map, b is the corresponding bias, and f is the activation function. The result of the convolution layer output goes to the downsampling layer, and down sampling is performed on each feature of the output in the convolution layer.

3. Proposed Method

Wavelet-decomposition vector c can be denoted as

$$c = [cA_n \ cD_n \ \dots \ cD_i \ \dots \ cD_1]. \quad (6)$$

Assuming that the length of the signal is L and the frequency is F_s , the highest frequency of the signal is $F_s/2$. The frequency range of the lowest layer cA_n is $(0, F_s/2^{n+1})$, with the size of $L/2^n$. The frequency range of cD_i is $(F_s/2^{i+1}, F_s/2^i)$, with the size of $L/2^i$. If we do STFT for c and select the window width as n_w , the sampling of cA_n is equivalent to the window width of about $2^{n-1} * n_w$ for the original signal, and the window width of cD_i is equivalent to that of the original signal $2^{i-1} * n_w$. In other words, if the frequency drops by one time, the window width increases by one time. Thus, we realize the effect of wavelet transform of large time windows at low frequency and small-time windows at high frequencies, almost without data redundancy.

Figure 3 shows the proposed deep-learning training. The predictor and target network signals are the magnitude spectra of the wavelet-decomposition vector of the noisy and clean audio signals, respectively. The network's output is the magnitude spectrum of the denoised signal. The regression network uses the predictor input to minimize the mean square error between its output and the input target. The denoised wavelet-decomposition vector is converted back to the time domain using the output magnitude spectrum and the phase of the noisy wavelet-decomposition vector. Then, the denoised speech can be obtained from the inverse wavelet transform.

4. Experiments and Discussion

The work used the Chinese Common Voice Corpus 6.1 subset of the Mozilla Common Voice dataset [17] to train and test our proposed method. Vehicle noise (Volvo) from the NOISEX-92 database [18] was taken as the noise source. The speech and noise were resampled at 16 kHz. The signal-to-noise ratios (SNR) of 5, 0, and -5 dB were set to compare the denoising effect.

Morse wavelet function was used in DWT. Another DNN method used STFT and convolution neural network for comparison [19]. The window length of 64 of STFT was adopted for our proposed method and those of 64 and 256 were adopted for the compared method. Hamming window with an overlap of 75% was used in all cases.

Figure 4 shows the clean speech and the noisy speech with different SNRs in the time domain and spectrogram.

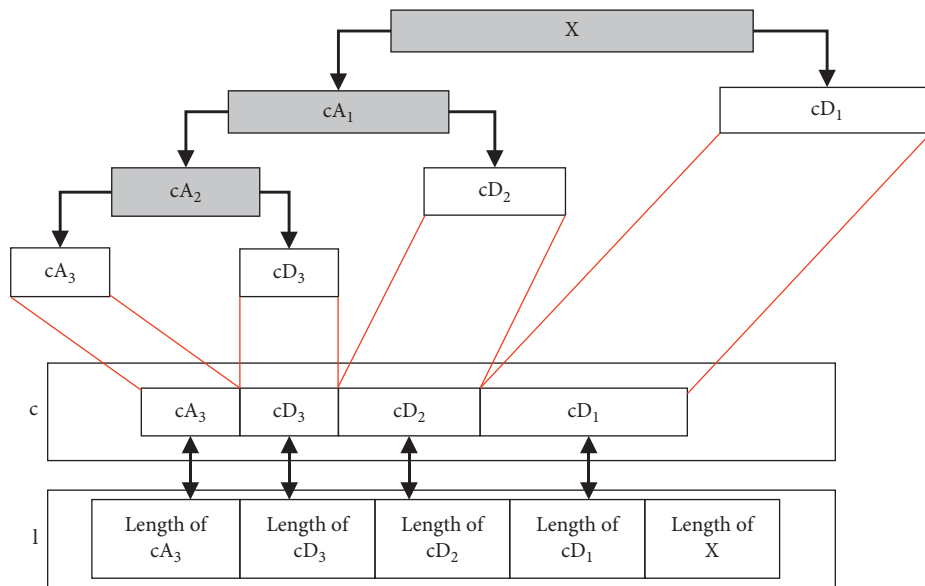


FIGURE 1: The diagram of wavelet decomposition.

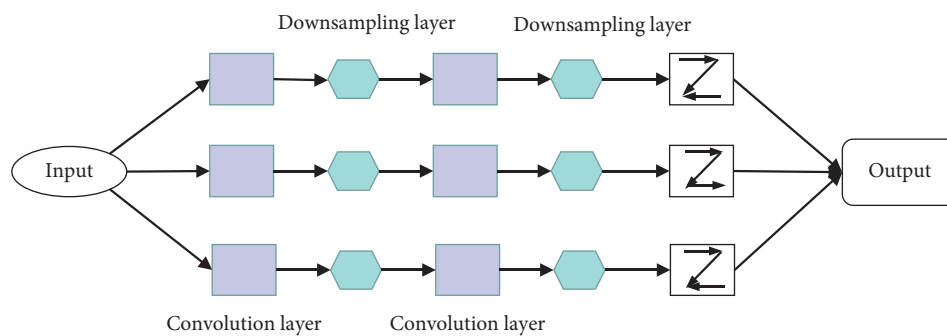


FIGURE 2: The diagram of the convolutional neural network model.

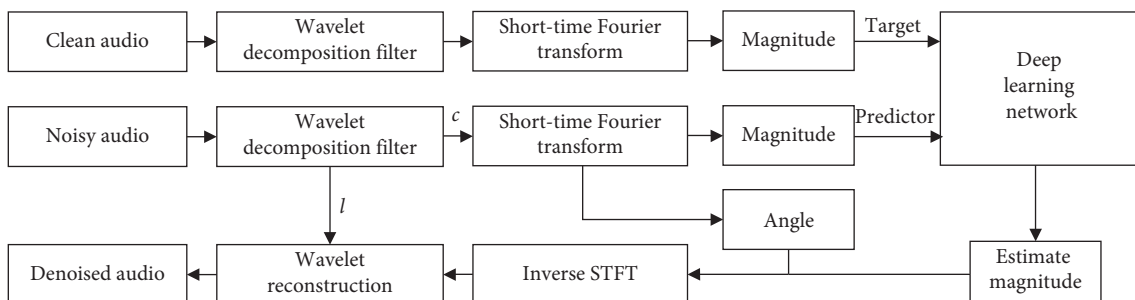


FIGURE 3: Block diagram of our proposed method.

The noise pollutes the noise in the broadband frequency. As the SNR decreases, more speech information is drowned out.

Figure 5 shows the speech signal enhanced by subtracting amplitude spectra. The noise has been reduced partly. The spectrogram shows that the rough points of the original noisy speech have been reduced to a large extent. Due to the half-wave rectification of negative values, small, independent peaks appear on the random frequency of the multiframe spectrum. Transformed to the time domain, these peaks sound like multiple vibratos with random

frequency changes between frames, which is commonly referred to as music noise.

In Figure 6, after Wiener filtering, the speech signal polluted by noise has been improved to a certain extent. However, there are still some noises after Wiener filtering, related to the filtering characteristics of the Wiener filter.

Figures 7–9 show the denoising results using the proposed method and the compared method, respectively. The results of the proposed method show a better denoising effect from high to low SNRs in the whole frequency range. The compared

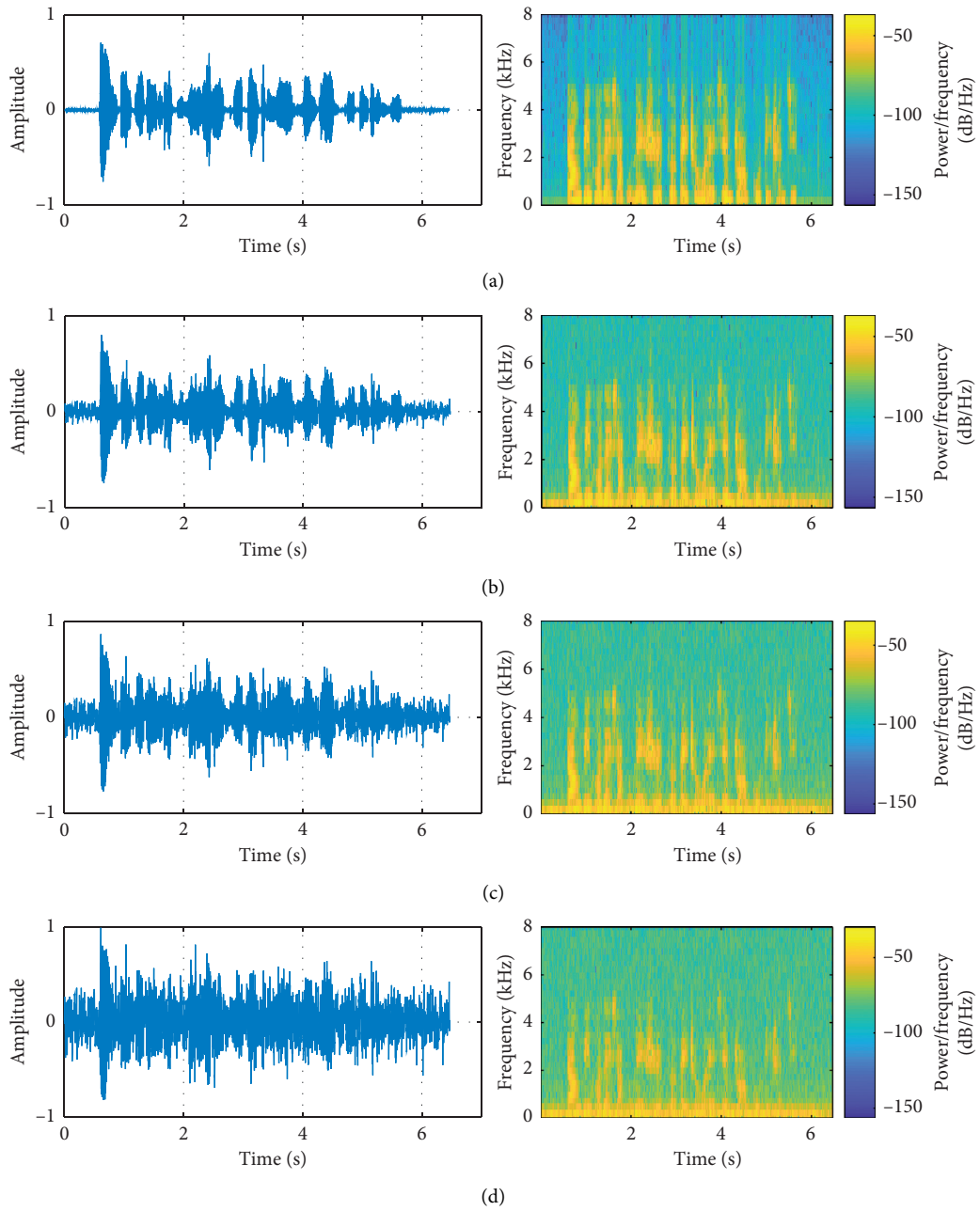


FIGURE 4: Clean speech and noise speech. (a) clean speech; (b) noisy speech (SNR 5 dB); (c) noisy speech (SNR 0 dB); (d) noisy speech (SNR-5 dB). Left: time domain. Right: spectrogram.

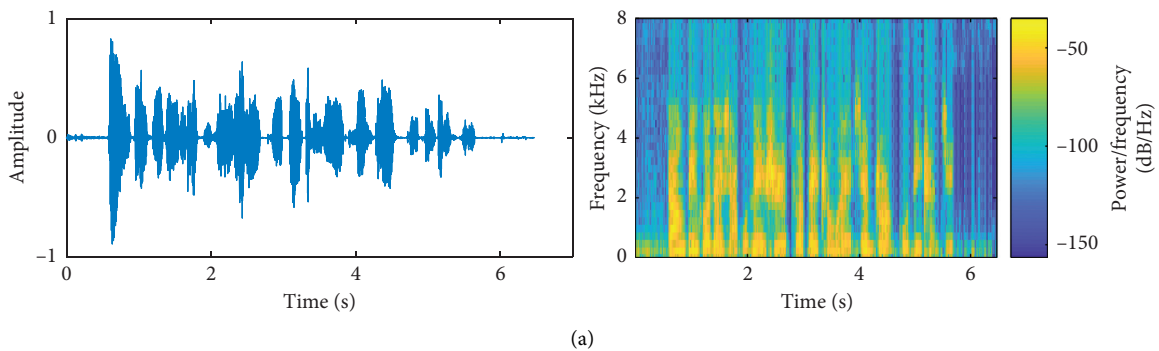


FIGURE 5: Continued.

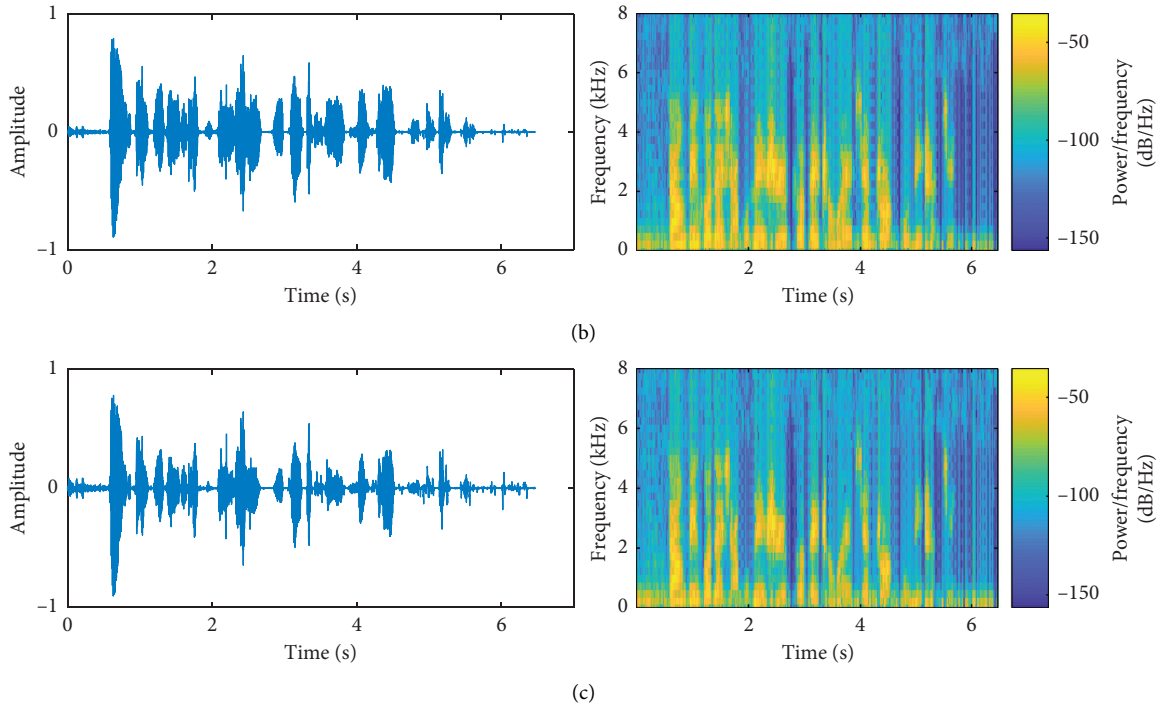


FIGURE 5: Enhanced speech using spectral subtraction. (a) SNR=5dB; (b) SNR=0 dB; (c) SNR=-5 dB. Left: time domain. Right: spectrogram.

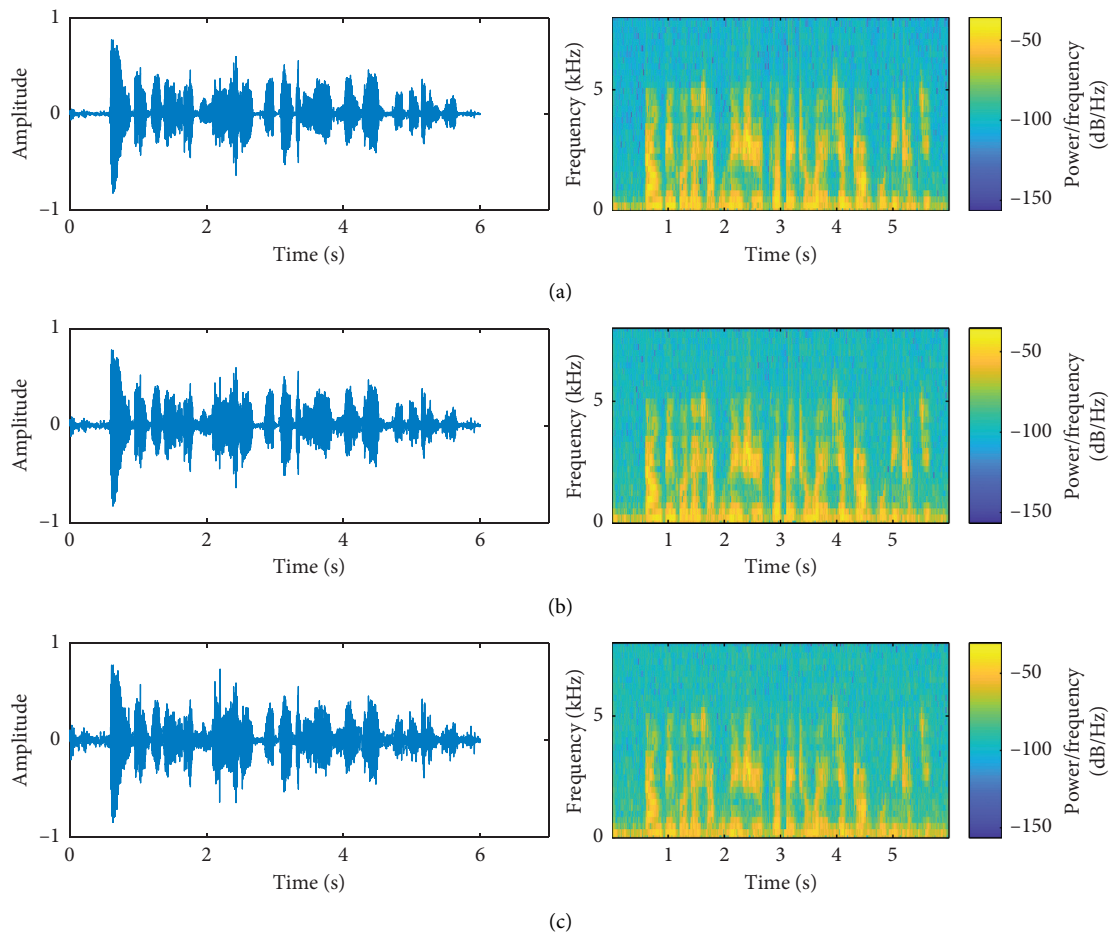


FIGURE 6: Enhanced speech using Wiener filtering. (a) SNR=5 dB; (b) SNR=0 dB; (c) SNR=-5 dB. Left: time domain. Right: spectrogram.

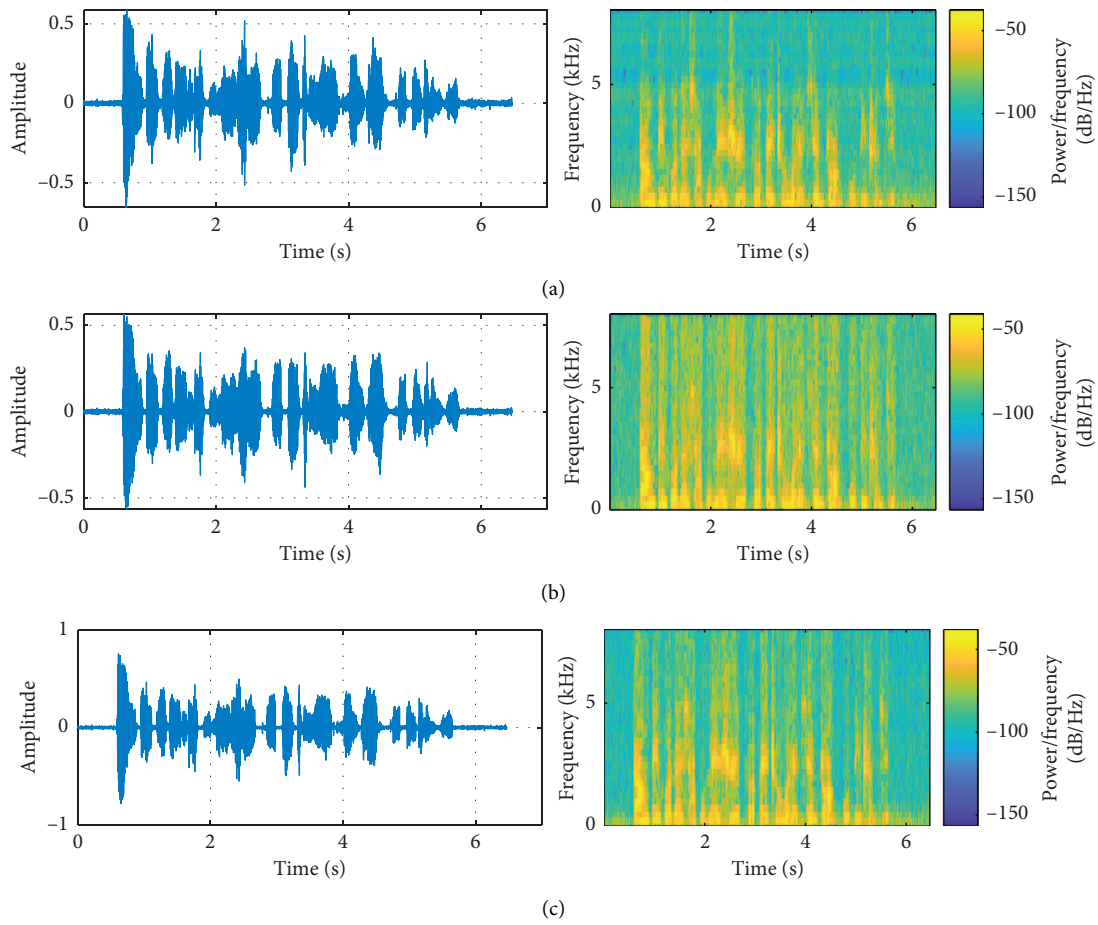


FIGURE 7: Enhanced speech (SNR=5 dB). (a) The proposed method. (b) The comparison method with 256 window lengths. (c) The comparison method with 64 window lengths. Left: time domain. Right: spectrogram.

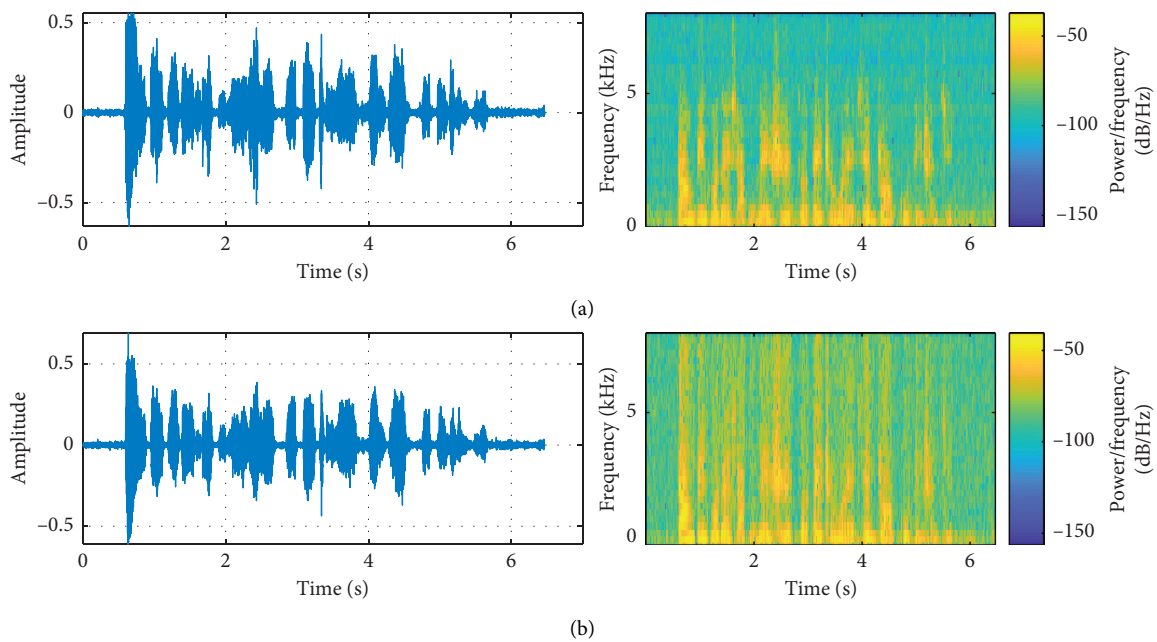


FIGURE 8: Continued.

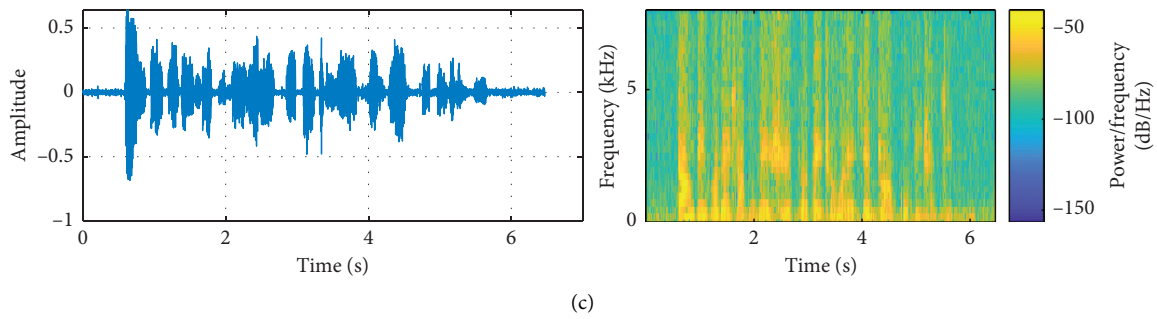


FIGURE 8: Enhanced speech (SNR=0 dB). (a) The proposed method. (b) The comparison method with 256 window lengths. (c) The comparison method with 64 window lengths. Left: time domain. Right: spectrogram.

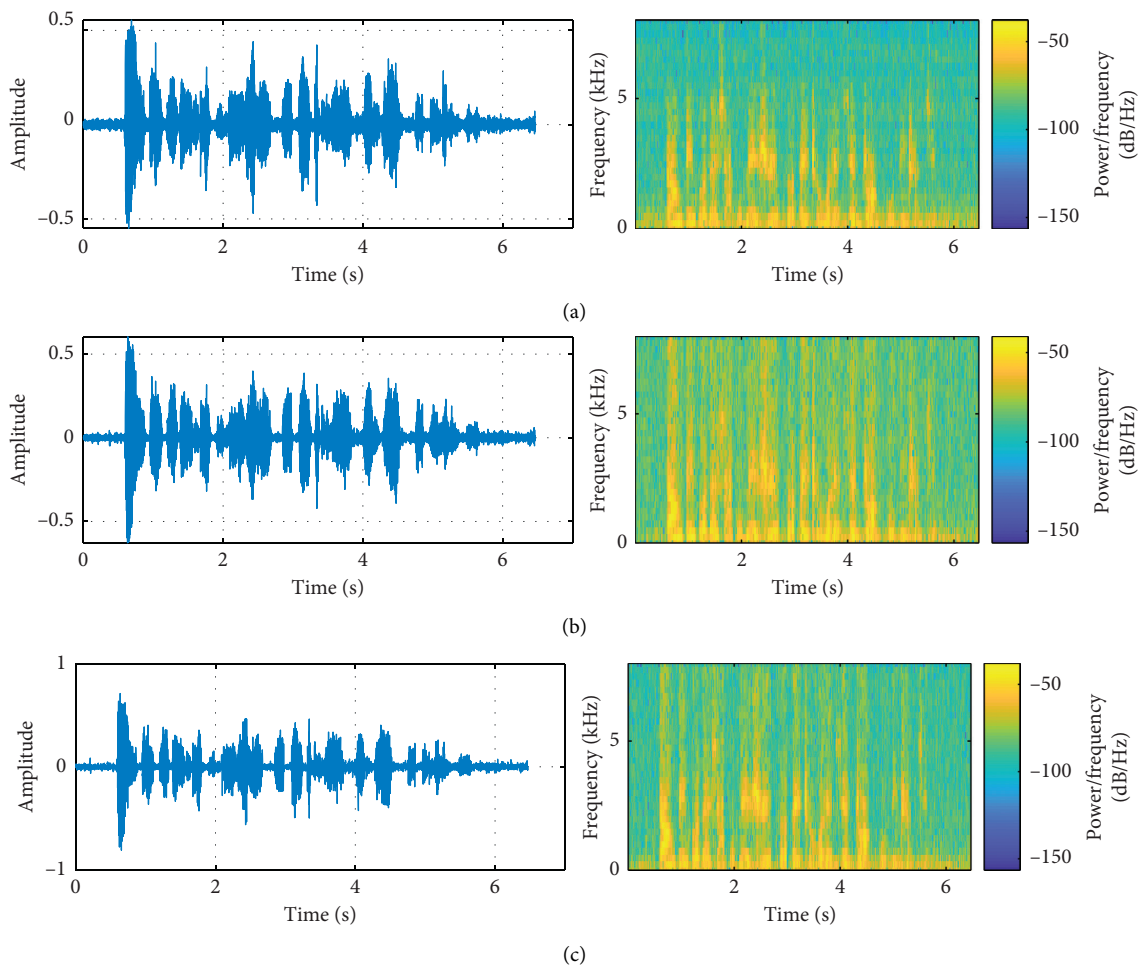


FIGURE 9: Enhanced speech (SNR=-5 dB). (a) The proposed method. (b) The comparison method with 256 window lengths. (c) The comparison method with 64 window lengths. Left: time domain. Right: spectrogram.

method with the window length of 256 achieves some noise reduction effect, but it performs poorly in the high-frequency band. The compared method with 64 window lengths performs some superiority in the high-frequency band, but is still inferior to the proposed method. In the process of the wavelet

transform, the signal energy in the frequency band remains the same with the reduced noise energy, which improves the SNR in the frequency band and denoising effect. Table 1 shows the SNR of the denoising speech, indicating the proposed method is an improvement of the compared method.

TABLE 1: SNR for denoising speech.

Noisy speech	Proposed method	Comparing method (256 window length)	Comparing method (64 window length)
5	16.7	15.5	12.6
0	14.3	13.7	12.3
-5	13.2	12.5	9.5

5. Conclusions

For the proposed method in the work, the predictor and the target network signals were the amplitude spectra of the wavelet-decomposition vector of the noisy audio signal and the clean audio signal, respectively. The output of the network was the amplitude spectrum of the denoising signal. The regression network used the input of the predictor to minimize the mean square error between its output and input targets. The denoised wavelet-decomposition vector was transformed back to the time domain using the output amplitude spectrum and the phase of the denoised wavelet-decomposition vector. Then, the denoised speech was obtained by the inverse wavelet transform.

The proposed method overcame the problem that the frequency and time resolution of STFT could not be adjusted. Besides, since the noise energy was gradually reduced during wavelet decomposition, the noise reduction effect of each frequency band was improved. The experimental results showed that the proposed method has a good denoising effect in the whole frequency band.

Data Availability

The datasets and codes of this paper for the simulation are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the project of Hubei University of Arts and Science (XK2020021), Natural Science Foundation of Guangxi (No. 2018GXNSFAA281276), and Liudong Science and Technology Project (20200108).

References

- [1] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 27, no. 2, pp. 113–120, 1979.
- [2] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 32, no. 6, pp. 1109–1121, 1984.
- [3] P. Scalart and J.V. Filho, "Speech enhancement based on a priori signal to noise estimation," in *Proceedings of the Acoustics, Speech, and Signal Processing, 1996 ICASSP-96*, pp. 629–632, IEEE, Atlanta, Georgia, USA, May 1996.
- [4] Y. Ephraim and H. L. Van Trees, "A signal subspace approach for speech enhancement," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 4, pp. 251–266, 1995.
- [5] K. W. Wilson, B. Raj, P. Smaragdis, and A. Divakaran, "Speech denoising using nonnegative matrix factorization with priors," in *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4029–4032, IEEE, Las Vegas, NV, USA, March 2008.
- [6] P. C. Loizou, *Speech Enhancement: Theory and Practice*, CRC Press, Florida, United States, 2007.
- [7] O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE/ACM Transactions on audio, speech, and language processing*, vol. 22, no. 10, pp. 1533–1545, 2014.
- [8] D. Liu, P. Smaragdis, and M. Kim, "Experiments on Deep Learning for Speech denoising," in *Proceedings of the Fifteenth Annual Conference of the International Speech Communication Association*, pp. 1–5, (ISCA), Singapore, September 2014.
- [9] Y. Wang, J. Du, L.-R. Dai, and C.-H. Lee, "A gender mixture detection approach to unsupervised single-channel speech separation based on deep neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 7, pp. 1535–1546, 2017.
- [10] H. Zhao, S. Zarar, I. Tashev, and C.-H. Lee, "Convolutional-recurrent neural networks for speech enhancement," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2401–2405, IEEE, Calgary, AB, Canada, April 2018.
- [11] K. Tan and D. Wang, "A convolutional recurrent neural network for real-time speech enhancement," in *Proceedings of the Proc. Interspeech*, pp. 3229–3233, Hyderabad, India, June 2018.
- [12] S.-W. Fu, Y. Tsao, X. Lu, and H. Kawai, "Raw waveform-based speech enhancement by fully convolutional networks," in *Proceedings of the 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 6–12, IEEE, Kuala Lumpur, Malaysia, December 2017.
- [13] I. Daubechies, "Where do wavelets come from? A personal point of view," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 510–513, 1992.
- [14] S. He, J. Chen, Z. Zhou, Y. Zi, Y. Wang, and X. Wang, "Multifractal entropy based adaptive multiwavelet construction and its application for mechanical compound-fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 76–77, pp. 742–758, 2016.
- [15] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 45, no. 5, pp. 485–560, 1992.
- [16] X. Ma, C. Zhou, and I. J. Kemp, "Automated wavelet selection and thresholding for PD detection," *IEEE Electrical Insulation Magazine*, vol. 18, no. 2, pp. 37–45, 2002.
- [17] <https://voice.mozilla.org/en>.

- [18] A. Varga and H. J. M. Steeneken, "Assessment for automatic speech recognition: II. NOISEX-92: a database and an experiment to study the effect of additive noise on speech recognition systems," *Speech Communication*, vol. 12, no. 3, pp. 247–251, 1993.
- [19] <https://ww2.mathworks.cn/help/deeplearning/ug/denoise-speech-using-deep-learning-networks.html>.

Research Article

A Study on the Relationship between Public Derivative Big Data and Industrial Policymaking: Taking Bike Sharing as an Example

Huilin Song 

School of International Economics and Trade, Jiangxi University of Finance and Economics, Nanchang 330013, China

Correspondence should be addressed to Huilin Song; songhuilin@jxufe.edu.cn

Received 6 May 2021; Accepted 3 July 2021; Published 16 July 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Huilin Song. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Smart government is an important part of the smart world. The use of big data analysis technology can effectively improve the government's ability of fine management. Taking China's bike-sharing industry as the research object, we study the relationship between public-derived big data and industrial policy. First, a feature-enhanced short text clustering method is proposed to perform topic clustering on publicly derived big data. Second, keyword extraction based on word frequency is used to quantify the text of industrial policy. Finally, time is taken as the main line to analyze the co-occurrence of clustering topics and keywords. The results show that (1) the feature enhancement method we proposed can effectively improve the clustering effect. (2) There is a great correlation between the industrial policy and the information mined by Weibo, but there is an obvious lag. Rational use of public-derived big data will effectively help the industrial policy to be released in a better and faster way.

1. Introduction

With the development of mobile communication technology, the mobile Internet, and the popularization of mobile smart devices, the characteristics of the widespread of mobile Internet information and the large coverage have made people enter an era of information explosion. Since the era of “We-Media,” the impact of public-generated big data [1] on public policy changes has become more pronounced [2]. Among the many ways of publishing information on the Internet, Weibo (Chinese microblog) has the characteristics of limiting the total number of input words (no more than 140 words), as well as strong interactivity and time-sensitive, makes the focus information in the text easier to be highlighted. It has gradually become an important information channel for the general public to discover news, release information, and discuss. In China, almost all mainstream official media and government propaganda departments at all levels have registered Weibo accounts. These official Weibo accounts of the government or media have high updating frequency and strong interaction, which have become a new platform for communication with the public and a channel for direct dialogue between the public and policymakers [3].

The sharing economy is based on the Internet as a platform to provide individuals or organizations with idle resources to temporarily transfer the right to use the service. Taking bike sharing as an example, it effectively improves the utilization efficiency of social resources, optimizes the allocation of resources, solves the “last kilometer” traffic pain point of residents, and meets the needs of different groups. However, driven by the brutal and disorderly capital, bicycle-sharing companies have the uncontrolled release of shared bicycles in cities to quickly seize the market, resulting in oversupply. This behavior not only violates the original intention of the “sharing economy” to optimize the allocation of resources but also brings new social problems. The disorderly parking of bicycles, the disposal of old bicycles, and other problems not only bring trouble to the city management but also affect the normal traffic order. Excessive competition leads to poor operation of enterprises, and some enterprises even delay refund or withhold user deposits. This series of problems exposed the lag in the government's policy formulation for the development of the new industrial norms of the bike-sharing industry. Many users of the above-mentioned problems have posted on Weibo for the first time, but due to the fragmented nature of Weibo information, it is not easy to collect, organize, and mine this

information. Therefore, how to mine the effective information of shared bicycles in Weibo, analyze its relationship with industrial policies, and finally provide decision support for the government to formulate corresponding industrial policies is the focus of this paper.

The overall research framework of this paper is shown in Figure 1. Firstly, relevant data are collected through web spider, then topic extraction and classification of Weibo data is carried out through the BTM topic model, keyword extraction based on word frequency is carried out on policy text, and finally, co-occurrence of problems and policies is analyzed based on time series, to provide strong decision support for the formulation of industrial policies.

2. Related Work

2.1. The Public Derivative Big Data and Public Policy. All social activities will generate data, and various technologies for analyzing massive data have emerged. How to use data and data analysis technology to optimize public policy issues and improve the quality of decision making has become a research hotspot. In the field of management, big data analysis technology is the first to be applied to the business field by large multinational companies [4]; through the mining of multidimensional data, it provides effective support for the final decision. With the application of this technology in public government affairs, research on the relationship between public management, public policy and public-derived big data has gradually emerged [6–9]. Public-derived big data refer to the data released by non-professionals around public topics and public affairs on the Internet platform. These data are characterized by complexity, diversity, and low-value density.

Cai and Yang [10] put forward an application framework in a big data environment that includes four parts: the construction of past case base, the detection and analysis of current social public opinion, the early warning of future public opinion, and the support of public policy decision making. Different computational models have been applied to the mining of commonly derived big data, including text mining, semantic understanding, sentiment analysis, and hot spot discovery [11]. Ma et al. [12] introduced the topic discovery model LDA in the analysis of public messages on the interactive platform of Chaoyang District, Beijing. It first measures the cost elements, then quantifies the service efficiency, and discusses the relationship between the public service efficiency and the cost. Chun et al. [13] used the Bayesian classification model to analyze the public environmental policies of different races under the condition that income and political behavior and other factors remained unchanged. Li et al. [14] improved the traditional MapReduce method and mined the relationship between text feature vectors and the development of public opinion.

2.2. The Topic Discovery Model for Short Texts. Text topic discovery has always been a focus in the field of natural language processing. The bag of words (BOW) model was first used in the discovery of text topics. The model regards

each document as a combination of multiple words and assumes that the relationship between words is independent [15]. Meanwhile, the default words in the document are out of order, ignoring the grammar and word order of the text of the document and simply using words to represent the topic of the document. This unordered method was gradually replaced by the LDA (latent Dirichlet allocation) [16] topic model due to poor performance. Based on Twitter, Weng et al. [17] considered combining a user's tweets into a single document and then used LDA for training. Similarly, Honey and Herring [18] took into account that the microblogs published by different people may involve many different topics, so they used Twitter messages containing the same words for aggregation. However, such a method greatly depends on the validity of the dataset, and the effect after aggregation cannot be guaranteed. After the BTM [19] model was proposed, it showed good performance not only in long text processing but also in short text processing and was considered to be a good substitute for the traditional topic model of LDA. Tang [20] used BTM model to represent microblog feature vectors and combined the traditional vector space model (VSM) with BTM according to certain weights to make up for the deficiency of VSM. Zhang et al. [21] studied the use of BTM's topic-vocabulary matrix as the external knowledge to expand the VSM vector, effectively solving the problem of microblog data sparseness.

3. Feature Reinforcement Topic Model for Short Text

3.1. Feature Vector Enhancement Method. The multimeaning or synonym of a word is the key to semantic understanding, especially in the task of text topic discovery in a specific domain. Therefore, in this task, the model performance is often poor because of the special definition of words in a certain domain. For example, the word "WeChat" (a mobile instant messaging app that can be used for payment) may appear for two different topics of social interaction and payment, so social interaction and payment are the potential topics of "WeChat," and we need to strengthen the features, respectively. The enhancement method adopted in this paper is VSM. Figure 2 shows the feature enhancement method.

Formally, for a document D that needs to discover a topic, it may contain feature f , which contains the following three cases:

- (1) When feature f only belongs to the topic T_i , T_i is used to identify feature f during reinforcement, and the weight of the topic T_i is denoted as the weight of feature f .
- (2) When the potential topic of feature f includes $T_{i_1}, T_{i_2}, \dots, T_{i_m}$, m topics need to identify the characteristics of the current document at the same time.
- (3) When feature f does not belong to any topic T , this feature is not reinforced, and f itself is used to identify the document feature.

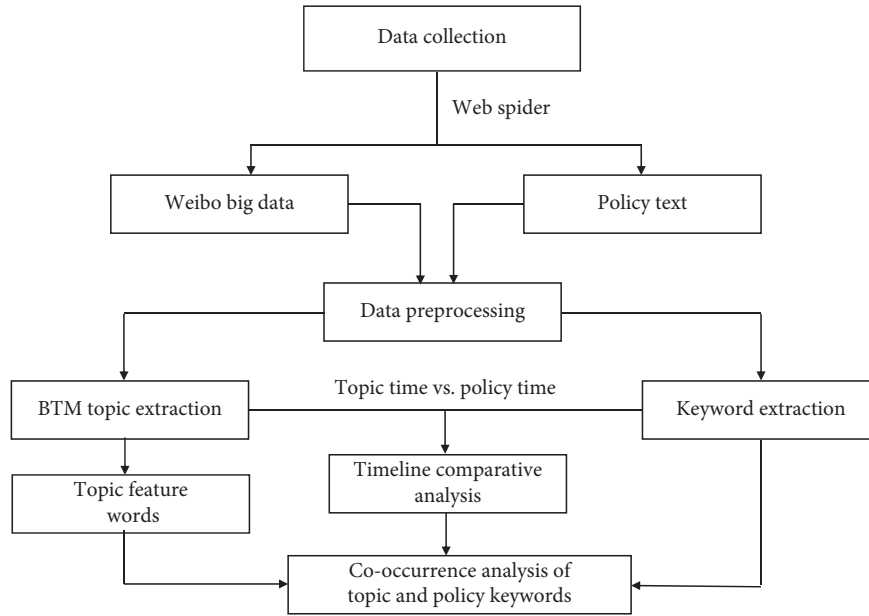
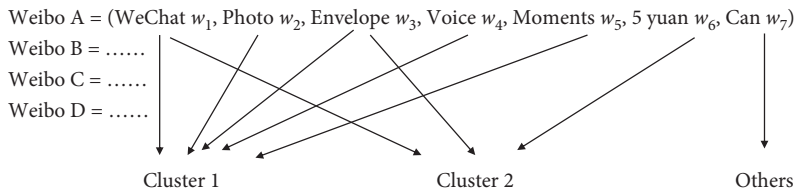


FIGURE 1: A framework of research on the relationship between public-derived big data and policymaking.

Word	WeChat	Photo	Envelope	Voice	Moments	5 yuan	Can
Weights	0.20	0.13	0.22	0.12	0.11	0.19	0.04

Vanilla VSM

Weibo feature vector representation:



Feature enhancement

Clustering results:

Cluster 1 (social) = {Weibo A, Weibo D,} description : WeChat, photos, envelopes, voice, moment.....

Cluster 2 (pay) = {Weibo B, Weibo C,} description : WeChat, envelopes, 5 yuan.....

Enhanced Weibo A = (feature1 : f_1 , feature2 : f_2 , Can : w_8)

Where $f_1 = a_1w_1 + w_2 + b_1w_3 + b_1w_4 + w_5$, $f_2 = a_2w_1 + b_2w_3 + c_2w_6$

- The coefficients a, b, and c are determined by the distribution probability of feature words under different topics

After feature enhancement

FIGURE 2: An example of feature enhancement.

For vector V , it is assumed that f_1, f_2, f_3 , and f_4 belong to topic T_1 , f_4 and f_5 belong to topic T_2 , and f_6 and f_7 do not belong to any topic. V' is the Weibo vector after feature enhancement. Figure 3 shows the process of feature enhancement.

There are three advantages after feature combination: (1) T_1 , the main potential topic of the text, is highlighted; (2) selective reinforcement also takes into account the possibility of f_4 becoming a secondary potential topic T_2 ; and (3) for general terms with no topic such as f_6 and f_7 ,

$$V = (f_1 : w_1, f_2 : w_2, f_3 : w_3, f_4 : w_4, f_5 : w_5, f_6 : w_6, f_7 : w_7)$$

$$V' = (T_1 : w_1 + w_2 + w_3 + w_4, T_2 : w_4 + w_5, f_6 : w_6, f_7 : w_7)$$

FIGURE 3: Formal process of feature enhancement of VSM vector.

their sparsity can be kept as suppression. The enhanced VSM vector, to some extent, reduces the dimension and highlights the main potential topic of the document.

When the fixed vector dimension M is maintained, the enhanced vector can retain more original information of the document, highlight the main features, and better describe the document.

3.2. BTM Model. BTM is a generation model, and its generation process is shown in Figure 4. The BTM model consists of three layers, namely, word pair, topic, and word, in which word pair to topic obeys Dirichlet distribution and topic to word obeys polynomial distribution. Considering that the entire document is a mixture of multiple topics, the BTM model algorithm can further alleviate the problem of data sparsity and facilitate the topic discovery of short texts by learning from the global topic library.

Specifically, the algorithm process of the BTM model can be described as follows.

Firstly, the word distribution under a single topic z is plotted, i.e., $\varphi_z \sim \text{Dir}(\beta)$, where $\text{Dir}(\beta)$ represents the Dirichlet distribution of words.

Secondly, the global topic distribution of the whole document is plotted, i.e., $\theta \sim \text{Dir}(\alpha)$, where $\text{Dir}(\alpha)$ represents the polynomial distribution of the topic.

Next, operate on b for each word, assuming $b = (w_i, w_j)$.

- (a) Select a topic z from the global topic distribution θ , i.e., $z \sim \text{Multi}(\theta)$.
- (b) Choose two words w_i and w_j from topic z , i.e., $w_i, w_j \sim \text{Multi}(\varphi_z)$.

According to the above process, the joint probability $P(b)$ of the word against b can be calculated by the following formula:

$$P(b) = \sum_z P(z)P(w_i|z)P(w_j|z) = \sum_z \theta_z \varphi_{i|z} \varphi_{j|z}. \quad (1)$$

Among them, $P(w_i|z)$, $P(w_j|z)$, respectively, represent the probability that words w_i and w_j belong to topic z . Further, the probability of the entire document topic is calculated by the following formula:

$$P(B) = \prod_{(i,j)} \sum_z \theta_z \varphi_{i|z} \varphi_{j|z}. \quad (2)$$

3.3. BTM Model Based on Feature Enhancement. The generation of VSM vectors by Weibo text requires the selection of appropriate terms as vector dimensions. The traditional document frequency method based on statistics can only remove noise through screening and cannot solve the problem of feature sparsity caused by the small number of words in Weibo. For this reason, we propose a feature enhancement method based on BTM, the process of which is shown in Figure 5.

In the process of BTM modeling, we construct the feature distribution matrix for the topic and then use the topic-lexical matrix to consolidate the vector features.

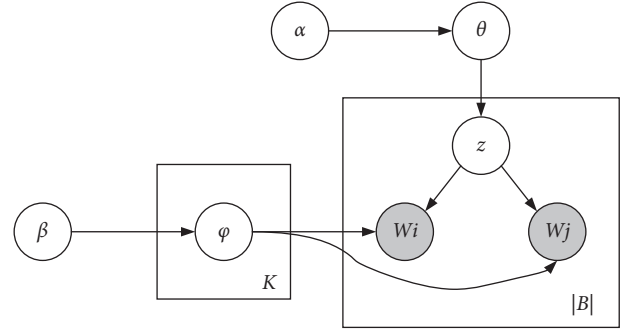


FIGURE 4: BTM model generation process diagram.

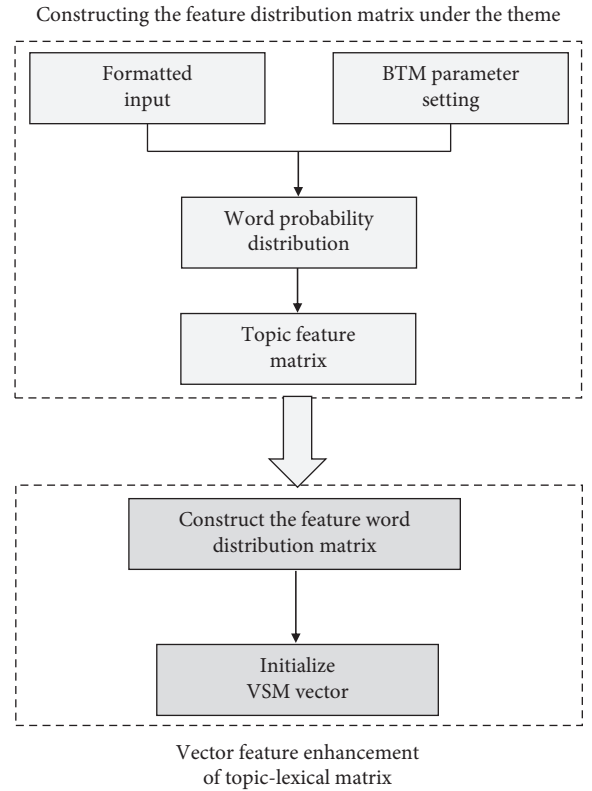


FIGURE 5: Flowchart of feature enhancement based on BTM.

3.3.1. Constructing the Feature Distribution Matrix under the Topic

Step 1: format the input. Input the VSM vector of each behavior in the aggregated document, and the VSM vector is generated after the selection of Weibo short text participles, stoppages, and features. First, we generate the dictionary text from all the words after segmentation, then map the VSM vector with the dictionary text, and finally get the formatted input text.

Step 2: parameter setting. In the BTM modeling in this paper, the parameter α is $50/K$, where K is the number of topics, and different values are set for different tasks. The

parameter β was 0.01; the number of iterations is set to 1000.

Step 3: generation of the word probability distribution. The words under each topic are arranged in descending order according to their distribution probability and spliced into the format of “words: weight,” which is stored in the probability text file. One line represents the distribution probability of all words under a topic.

Step 4: generate the topic-feature matrix. The words with a probability higher than 0.001 in the probability text were selected as candidate words and used as the feature words of the topic, and ID mapping was carried out to generate the topic-feature matrix.

3.3.2. Vector Feature Enhancement of Topic-Lexical Matrix.

After the topic-feature database is obtained, to facilitate the calculation of the algorithm, we need to reverse calculate the probability distribution of each candidate feature under different topics and carry out normalization processing, to allocate the weight during selective reinforcement. The matrix structure is as follows:

$$\begin{bmatrix} P_{00} & P_{01} & \cdots & P_{0k} \\ P_{10} & P_{11} & \cdots & P_{1k} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n0} & P_{n1} & \cdots & P_{nk} \end{bmatrix}, \quad (3)$$

where p_{ij} represents the probability of the i -th word under topic j , i is 0 to n , j is 0 to k , n is the number of candidate feature words, and k is the number of candidate topics. Then, the VSM vector is initialized. In this paper, each weight vector is added k dimensions and initialized to 0, that is, $t_i = 0$. According to the feature vector enhancement method introduced in Section 3.1, for the vector V_{weight} , search in the feature word distribution matrix according to each word in the VSM vector, complete the feature enhancement process of all topics, and update the weight value in V_{weight} .

4. Empirical Data

4.1. Data Collection. The data time interval selected in this paper is from July 2016 to October 2017, which covers the three market development stages experienced since the birth of shared bikes, which can depict the development trajectory of shared bikes. Figure 6 shows the bike-sharing market and user timeline including three stages of development. The period of the first stage is before October 2016. At this stage, the market is mainly concentrated in developed cities such as Beijing, Shanghai, Guangzhou, and Shenzhen. There are no more than five service providers and the user scale is just over one million. The period of the second stage is from October 2016 to June 2017. At this stage, the market has covered the mainstream first-tier cities in China. The number of service providers has soared to dozens, and the user scale has reached about 50 million after experiencing explosive growth. The period of the third stage is after June 2017. At this stage, the market has begun to sink to third-tier cities. Many service providers have closed down or are forced to merge, and the user scale is stable

at about 70 million. To accurately reflect the characteristics of each stage, we named the first stage as the user training stage, the second stage as the demand outbreak stage, and the third stage as the elimination stage.

The data are divided into two parts: Weibo data and policy documents. Since the amount of Weibo data is huge, we use the multithreaded web spiders [22] based on the Scrapy framework for automatic collection. Table 1 shows the details of Weibo data acquisition, including data sources, time range, acquisition tools, keywords, and screening rules.

Due to the limited number of policy documents, we manually downloaded and collected them from the government website.

4.2. Data Preprocessing. Both Weibo and policy documents are in text format, which cannot be directly modeled and analyzed. Therefore, we preprocessed the text data first. Weibo data preprocessing is divided into three steps.

The first step is word segmentation. The word segmentation tool selected in this paper is Jieba word segmentation [23], which provides a custom user dictionary function. To achieve a better word segmentation effect in specific fields, the field words in the Sogou input method lexicon are selected in this paper to improve the effect of high score words.

The second step is to remove emoticons and stop words. First, emoticons are removed from all texts. Then, more than 3,000 discontinued words, including punctuation marks, meaningless numbers, mood particles, appellation words, etc., are selected from the online collection.

The third step is to calculate the weight and calculate the TF-IDF value according to the above-obtained text to form the original VSM vector after preliminary processing.

The policy document data preprocessing is relatively simple, only requiring word segmentation. The word segmentation method is similar to the first step of Weibo data preprocessing, and the method will not be described in detail.

5. Result and Analysis

5.1. Feature Enhancement Results. The initial VSM vector is feature-enhanced using the matrix of potential topic high-frequency words, that is, FE-VSM (feature-enhanced vector space model). In order to verify the effect, the following three kinds of vectors were used as references to conduct comparative experiments: (1) VSM vector without feature enhancement; (2) LDA topic vector; (3) feature vector combined with BTM and VSM. We compare the quality of clustering when the vector dimensions are 100, 300, 500, 700, 900, 1100, 1300, and 1500. The text similarity uses cosine similarity to perform standard K-means clustering. The k value is 12 (BTM optimal number of topics), and the initial center is randomly selected. To eliminate the chance of random selection, the results are averaged 10 times. The experimental results are shown in Table 2.

The results show that the effect of traditional VSM is the worst, and the effect of LDA is not good, and it is also very unstable, which is caused by the inadequacy of LDA in a short text. Combining the vector represented by the BTM

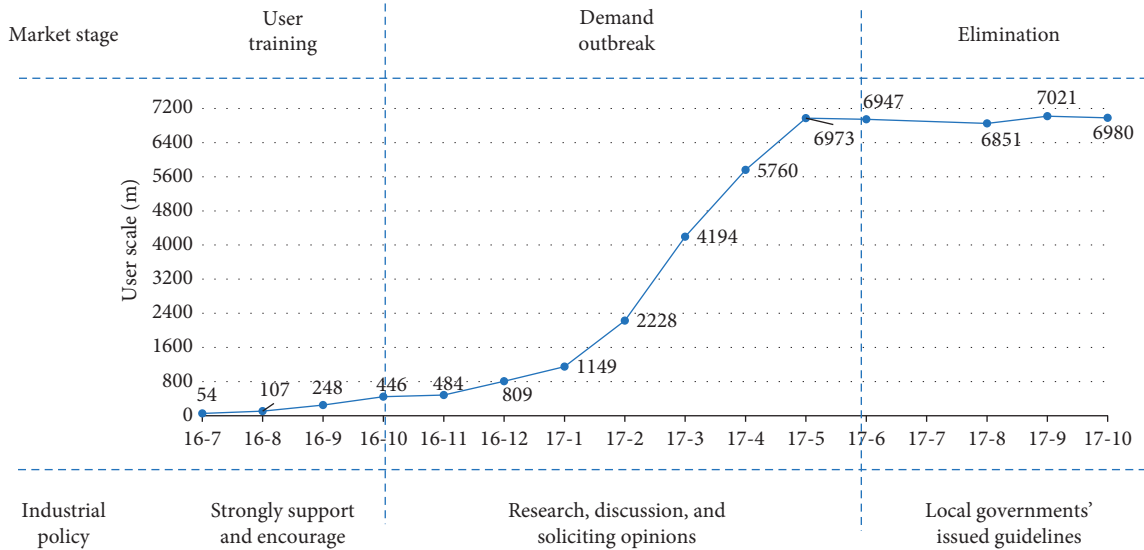


FIGURE 6: Timeline of the bike-sharing market and user size.

TABLE 1: Weibo posts' data collection details.

Source	http://www.weibo.com
Time range	From July 2016 to October 2017
Method	Multithreaded web spiders based on Scrapy framework Direct keywords: "shared bike," "Internet bike," etc.
Keyword	Bike brands: Mobike, Ofo, etc. Brand name: "little yellow car," "little blue car," and so on
Filtering rules	Rule 1: keep only original Weibo and exclude reposts Rule 2: keep only one Weibo for repeated Weibo

TABLE 2: Comparison of results of different topic clustering methods.

Vector dimension	VSM	LDA	BTM + VSM	FE-VSM
100	0.150612	0.18651	0.304817	0.244714
300	0.253034	0.185977	0.321333	0.365615
500	0.253241	0.211059	0.385545	0.435322
700	0.217615	0.266491	0.386406	0.355565
900	0.224406	0.150216	0.38029	0.416342
1100	0.227279	0.294109	0.373178	0.383586
1300	0.240335	0.208641	0.364237	0.395509
1500	0.224746	0.261475	0.369184	0.366157

The bold values indicate best results.

and the VSM vector according to a certain weight, the effect has been significantly improved. The clustering quality of BTM-enhanced vectors is also significantly better than that of unenhanced vectors, indicating that the method in this article solves the problem of data sparsity and expression diversity to a certain extent, and the enhanced vector can better describe Weibo text information.

5.2. User Training Stage Analysis

5.2.1. Weibo Topic Mining and Analysis. At this stage, the number of Weibo posts was relatively small, including 2,253 Weibo. The topic-feature word distribution results of topic

TABLE 3: Mining topic of Weibo in user education stage: distribution of feature words (excerpt).

Topic 1	Topic 2	Topic 3	Topic 4
Convenient	Clock in	One kilometer	Password
0.10314	0.03843	0.00942	0.00133
Give a like	Encounter	Green	Crack
0.08135	0.02414	0.00828	0.00128
Beautiful	Incredibly	Travel	Free
0.04432	0.01963	0.00764	0.00105
Like	Registered	Profit model	Hidden trouble
0.04012	0.01852	0.00732	0.00093
Big love	Expectation	Order	Deficit
0.34861	0.01284	0.00531	0.00081
Awesome	Ofo	Effectiveness	Locked
0.30234	0.00987	0.00397	0.00080
Fashion	Mobike	Health	Remove the seat
0.2396	0.00896	0.00381	0.00075
Cost-effective	App	Invention	Car chain
0.21753	0.00481	0.00362	0.00063

clustering are shown in Table 3. All Weibo are clustered into 4 different topics. Through the analysis of the feature words of topics 1, 2, 3, and 4, it can be found that most of the Weibo clustered by topic 1 is praise content released by users after using shared bikes. We call such topics advantage feedback. Topic-clustering Weibo expresses users' expectations and freshness, including the content of clocking in and

registering for the first time to use shared bikes. We call such topics public expectations. The Weibo with the topic tricluster is mostly the comments of media or opinion leaders on shared bikes. Such Weibo has a large number of words and is organized. We call this topic positive comments. Topic 4 of Weibo clustering contains a large number of characteristic words related to unlocking, password, and car lock, and the overall sentiment is relatively negative. This is because of cost control, technical constraints, and other factors. The mechanical locks used in the first-generation shared bicycles are easily cracked by users. At this time, many Weibo posts released the content of cracking the mechanical locks and riding for free. We classify such topics as having problems.

Topics 1, 2, and 3 show people’s positive emotions towards shared bikes from different perspectives, while topic 4 shows people’s negative emotions towards shared bikes. As shown in Figure 7, the proportion of microblogs with a different topic in the overall microblogs is as high as 92% with a positive attitude, which shows the public’s love for shared bikes and tolerance of its shortcomings in this period.

5.2.2. Keyword Extraction and Analysis of Policy Text. During this period, there were no policies for bike sharing. The only two policy documents, Guiding Opinions on Promoting Green Consumption and National Fitness Plan (2016–2020), only included policies related to bike sharing to encourage cycling and low-carbon travel. The results show that words such as “the whole people,” “fitness,” “green,” and “consumption” appear frequently. By looking at the sentences of these words, they are all policies guiding the whole industry. Therefore, it can be judged that in the early stage of bike-sharing development, there was no corresponding industrial policy support.

5.3. Demand Outbreak Stage Analysis

5.3.1. Weibo Topic Mining and Analysis. During this period, the number of microblogs increased sharply, including 31,259 Weibo. The topic-feature word distribution results of topic clustering are shown in Table 4. All microblogs are clustered into three different topics. Compared with the previous stage, the number of topics in the cluster decreased by one. This is because with the rapid expansion of bike-sharing enterprises, most cities have been covered, so the number of microblogs expecting the registration of bike-sharing enterprises decreased greatly. To be specific, the first topic is to express the convenience of daily use of shared bikes. The number of words on Weibo on this topic is relatively small, usually within a dozen words. Topic 2 is still the long microblog of media microblog number and opinion leaders’ analysis and evaluation of shared bikes. It can be seen from the table that the wording of these Weibo is a formal and overall affirmation of shared bikes. Topic 3 is similar to topic 1, except that this topic expresses the problems encountered in daily use and makes fun of them.

The first topic of daily use and the second topic of positive comments can still be regarded as people’s positive emotions towards shared bikes, while the third topic shows

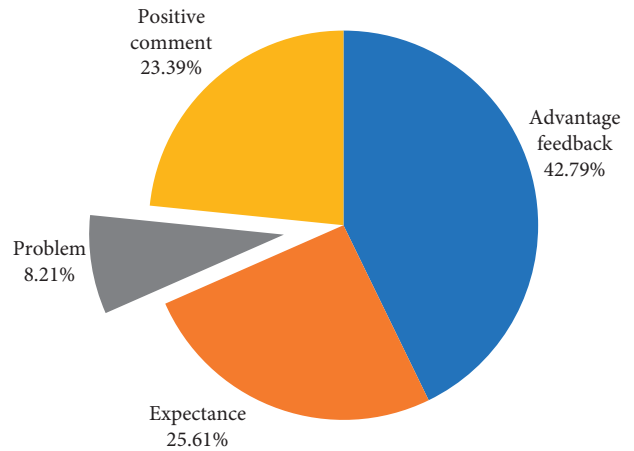


FIGURE 7: The proportion of the number of Weibo posts with different topics in the user training stage.

TABLE 4: Mining topic of Weibo in demand outbreak stage: distribution of feature words (excerpt).

Topic 1	Topic 2	Topic 3
Convenient	Short distance	QR code
0.15323	0.02071	0.10877
Free	Invention	Random parking
0.10334	0.01534	0.10060
Metro station	Pay	Locked privately
0.08641	0.01142	0.08265
Easily	Positioning	Deposit
0.06678	0.01087	0.06731
Reservation	Travel	Safe
0.06012	0.00881	0.05123
Praise	Lane	Damage
0.05621	0.00706	0.03876
Recommend	IPO	Broken
0.04130	0.00654	0.01134
Clock in	Quality	Phishing enforcement
0.02045	0.00586	0.00818

people’s negative emotions towards shared bikes. As shown in Figure 8, compared with the previous stage, due to the sharp increase in the number of users and the number of bicycles, the corresponding service quality of enterprises, and the lack of government supervision in this period, the proportion of the problematic topics rose from 8.21% to 31.3%.

5.3.2. Keyword Extraction and Analysis of Policy Text. Policy documents directly related to shared bicycles have already appeared at this stage. After removing “Internet bicycles,” “shared bicycles,” “relevant departments,” “operating companies,” “vehicles,” “transportation,” and other words with high frequency but no practical meaning, the final keywords can be divided into two parts: one is the description of policy planning, such as “travel,” “construction,” and “standard,” and the other is the description of specific policies, such as “parking” and “information”. It also includes specific policy descriptions of “parking” and “information” related to practical issues of shared bikes.

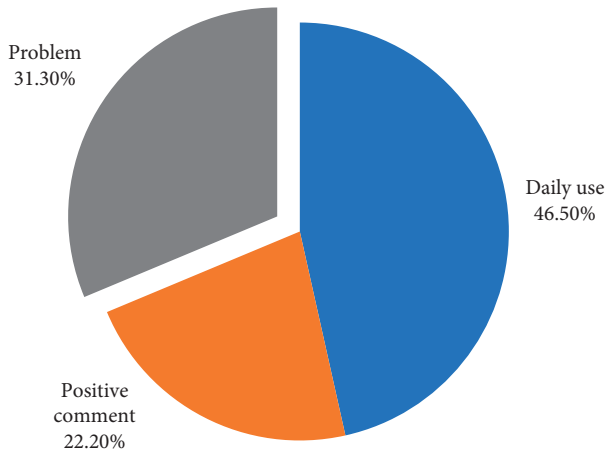


FIGURE 8: The proportion of the number of Weibo posts with different topics in the demand outbreak stage.

5.3.3. Comparison between Weibo Topics and Policy Texts.

First, we compare the collinearity between the topic words and the high-frequency keywords. The top five topic keywords in the three probability of the topic are “QR code,” “parking,” “locking,” “deposit,” and “safety.” The Weibo corresponding to these words concentrate on “QR code modification,” “parking and placing,” “locking without permission,” “deposit trouble,” and “driving safety.” Policy highest frequency in the text of the five words: “travel,” “park,” “construction,” and the “norm” and “information.” These words reflect the government in policy-making for travel “convenience,” “disorderly parking place,” “service construction,” “enterprise and individual behavior norms,” and “personal user information security” issues such as the focus. Among them, some problems in Weibo topics and policy text exist at the same time, such as the simultaneous occurrence of keywords: “stop the place,” and “locked” and “deposit” in the Weibo topic and issues such as policy text of “enterprise and individual behavior norms,” “security” and “supporting service construction” proposed by the corresponding lanes. Although “two-dimensional code” can also be classified as the category of “personal code of conduct,” but through the analysis of the original microblog, the problem of “two-dimensional code” described in the microblog contains the altered “two-dimensional code” fraudulent behavior. By comparison, the overall normative content of the policy is consistent with the problems excavated by Weibo, which indicates that the content of the policy formulation is in line with the basic demands of the public. Meanwhile, the absence of the problems represented by the high-probability feature word “QR code” in the policy text also reflects the lack of comprehensiveness of the policy text to some extent.

Second, we compared the relationship between the topic of the Weibo issue and the timeline of policy introduction. According to the statistics of the number of microblogs in the third topic every month, Figure 9 shows the change of the number of microblogs in the third topic in the eight months at this stage. It can be seen that the number of microblogs in January 2017 surged, which increased by more than four times compared with that in December 2016, and continued to increase in the following months. The first local bike-sharing

encouragement and regulation policy was issued in Chengdu in March 2017, and the first national draft for soliciting opinions was issued in May. The time point for various regions to introduce policies on a large scale has already come from August to October. For the same problem, the time for the government to issue policies has a certain lag compared with the time for Weibo topics. Taking the national policy issued by the Ministry of Transport as the time node, the lag period is 4 months, and for the time node of the local policy issued by Chengdu, the lag period is 2 months. Obviously, for the emerging “Internet+” industry of bike sharing, which takes no more than 18 months to sprout, develop, and stabilize, a period of 2–4 months will leave a regulatory gap for the development of the industry, which will not only affect the user experience but also affect the development of enterprises and even the whole industry.

5.4. Elimination Stage Analysis

5.4.1. Weibo Topic Mining and Analysis.

At this stage, the number of microblogs tended to stabilize and began to decline slowly, with a total of 27,627 microblogs. The topic-feature word distribution results of the Weibo topic clustering are shown in Table 5. 31,259 Weibo posts are clustered into five different topics. Compared with the previous stage, the number of topics in the clustering increased by two. The first topic is still to express the convenience of daily use of shared bikes, but its proportion is significantly reduced. This indicates that after a period of use, shared bikes have become a daily transportation tool for users, and users’ interest in Weibo has decreased. Both topic 2 and topic 3 are comments on shared bikes. The difference is that the sentiment of topic 2 is optimistic, while the microblogs of topic 3 are mostly objective and neutral comments. The fourth topic is the problems encountered by users in daily use and ridicule. Most of the microblogs in topic 5 focus on the description of the deposit problem of shared bikes.

As shown in Figure 10, compared with the previous stage, the proportion of negative emotions is still on the rise. On the one hand, this is due to the decrease of daily clocked microblogs; on the other hand, the number of policies in the previous stage is small and the specificity is not enough, and the effect of regulation is limited. Notably, a new topic (topic 5) highlights the issue of deposits. Through checking the original microblog corresponding to topic 5, we find that the deposit problem mainly focuses on two aspects: one is that users need to pay multiple deposits when using different brands of shared bikes; the other is that some bike companies cannot withdraw deposits on time due to their blind expansion and lack of funds. Therefore, both topic 4 and topic 5 can be regarded as the existing problem topics of Weibo mining.

5.4.2. Keyword Extraction and Analysis of Policy Text.

The proof policy of this stage country level already came on stage. From August to October, major cities rolled out local management policies for bike-sharing. After keyword extraction is removed, the high-frequency keywords include “positioning,” “special account,” “technology,” and other words, indicating that the policies at this stage are more specific than those at the previous stage.

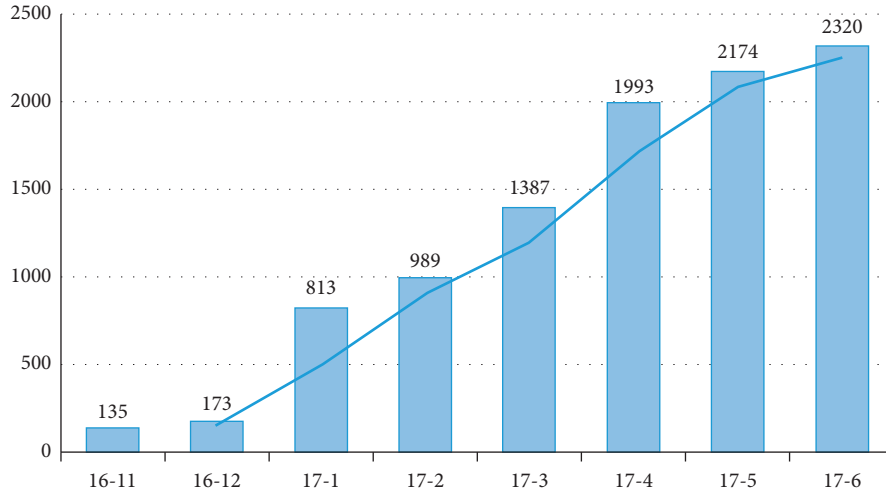


FIGURE 9: Monthly statistics on the number of Weibo posts of the third topic in the demand outbreak stage.

TABLE 5: Mining topic of Weibo in elimination stage: distribution of feature words (excerpt).

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Convenient 0.05452	Green 0.02071	System 0.00531	Random parking 0.01289	Deposit 0.06432
Free 0.04613	Travel 0.01534	Supervision 0.00519	Damage 0.1065	Zhima integral 0.0522
Applet 0.04171	Invention 0.01142	Regulatory 0.00408	Electric fence 0.0973	Account 0.04231
Easily 0.03983	Severe 0.01087	Protect 0.00391	QR code 0.0834	Extract 0.0387
Work 0.03286	Export 0.00881	12 years old 0.00353	Manned 0.0644	Rights 0.02231
Habit 0.02876	Information 0.00706	Real name 0.00298	Abandoned 0.0570	Credibility 0.02108
Campus 0.01971	Domestic 0.00654	Deficit 0.00223	Recycle 0.0445	Capital 0.01976
Shuttle 0.00886	Scenery 0.00586	Matching 0.00187	Disclosure 0.0386	Escape 0.00821

5.4.3. Comparison between Weibo Topics and Policy Texts.

First, we compare the collinearity between the topic words and the high-frequency keywords. The top five topic words for topic 4 and topic 5 are “parking,” “damage,” “electronic fence,” “deposit,” and “Zhima points.” The top five most frequently used words in policy texts were “positioning,” “self-regulation,” “special account,” “illegal parking,” and “technology.” By comparison, it is found that “parking in disorder” and “illegal parking” are the same, “electronic fence” and “positioning” is basically the same, and “deposit” and “ant integral” and “special account” are all about the description of user deposit. Therefore, the content of the policy at this stage is consistent with the theme of microblog mining, which indicates that the content of the policy meets the basic demands of the public. To sum up, the above empirical evidence shows that the public online opinions represented by Weibo posts can have a certain influence on industrial policy formulation.

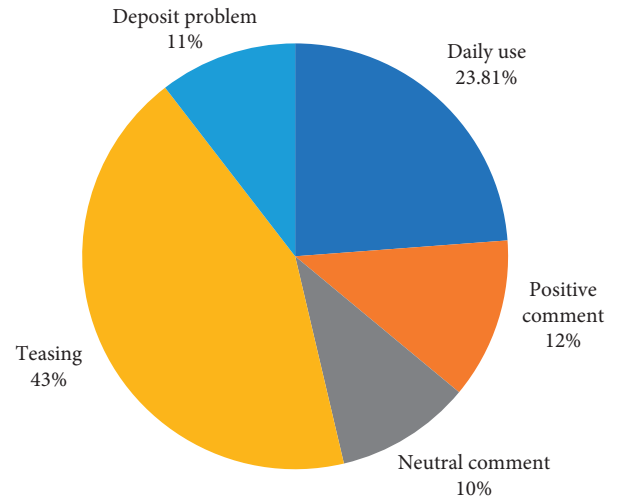


FIGURE 10: The proportion of the number of Weibo posts with different topics in the elimination stage.

Second, through the comparison of the timeline, we find that there is still a deviation between the time of the introduction of industrial policy and the peak time of network public opinion at this stage, with an average lag of 2 months. However, compared with the previous stage, this time has been reduced by half, which is related to the growth of the whole industrial chain and the tracking reports of other media (print media, TV media, and so on). The information released by users on Weibo has been spread and developed and has been concerned by the whole society.

6. Conclusion

First, a feature-enhanced short text clustering method is proposed to perform topic clustering on publicly derived big data. Second, keyword extraction based on word frequency is used to quantify the text of industrial policy. Finally, time is taken as the main line to analyze the co-occurrence of clustering topics and keywords. Through three different

stages of Weibo data mining and keyword analysis of policy text, we find the following.

In the stage of user training, bike sharing, as a typical “Internet+” industry, developed and grew in a short period by accurately solving the “pain points” of users and the power of capital. At this stage, more than 90% of microblogs expressed positive emotions for bike sharing, but at this stage, there was no effective industrial policy support.

In the stage of demand outbreak, with the increase of users, the fierce competition of manufacturers, the loss of bicycles, and the lack of supervision, the negative sentiment on Weibo at this stage accounted for 31.3%, which was nearly four times higher than the negative sentiment on Weibo at the previous stage, which was only 8%. Through the analysis of the timeline, we found that the time when the policy was issued had a certain lag compared with the problem time of public derivative big data mining represented by Weibo, and the regulatory vacuum caused by this lag led to the accumulation of negative emotions, which further affected the healthy development of the whole shared bike. Through the co-occurrence analysis of topic keywords and policy keywords, it is found that although the policy can cover most of the existing problems, it still omits hot issues like “QR code” in Weibo.

Finally, through the research of this paper, we find that the real situation of an industry can be reflected by collecting and clustering the information of industry on Weibo, and these reactions are highly correlated with the industrial policies of the industry. At the same time, we find that the formulation time of industrial policy is lagging behind that of the outbreak time of microblog information. For example, when it comes to industrial policy, proper consideration of the big data of network media represented by Weibo post will be conducive to the rapid introduction of industrial policy.

Data Availability

The data in this paper are all from the open information on the Internet, including two parts: the first part is the microblog data, which can be obtained from the website <http://www.weibo.com>; the second part is industrial policy documents from different government websites.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by the Science and Technology Research Project of Jiangxi Provincial Department of Education (grant no. GJJ200318) and the International Social Science Fund General Project (grant no. 16BJY082).

References

- [1] N. Zhang, “Analyzing public generated big data and restructuring government decision making process: review and prospect,” *Chinese Public Administration*, vol. 10, pp. 19–24, 2015.
- [2] B. Yu and K. Yang, “The interactive policy agenda-setting model in China’s Internet incidents: an empirical study on events concerning social justice,” *Journal of Nanjing Normal University (Social Science Edition)*, vol. 5, pp. 13–20, 2013.
- [3] Z. Deng and Q. Meng, “We-media agenda setting: a new path to formation of public policy,” *Journal of Public Management*, vol. 13, no. 2, pp. 14–22, 2016.
- [4] H. Chen, R. H. L. Chiang, and V. C. Storey, *Business Intelligence and Analytics: From Big Data to Big Impact*, Society for Information Management and The Management Information Systems Research Center, South Minneapolis, MN, USA, 2012.
- [5] N. Zhang, “Big data analysis of public derivatives and reconstruction of government decision-making process: theoretical evolution and research prospects,” *Chinese Public Administration*, vol. 10, pp. 19–24, 2015.
- [6] E. Johnston and Y. Kim, “Introduction to the special issue on policy informatics,” *The Innovation Journal: The Public Sector Innovation Journal*, vol. 16, no. 1, pp. 1–4, 2011.
- [7] E. G. Martin, R. H. MacDonald, L. C. Smith et al., “Policy modeling to support administrative decisionmaking on the New York State HIV testing law,” *Journal of Policy Analysis and Management*, vol. 34, no. 2, pp. 403–423, 2015.
- [8] M. I. Sirer, S. Maroulis, R. Guimerà, U. Wilensky, and L. A. N. Amaral, “The currents beneath the “rising tide” of school choice: an analysis of student enrollment flows in the Chicago public schools,” *Journal of Policy Analysis and Management*, vol. 34, no. 2, pp. 358–377, 2015.
- [9] K. A. Frank, W. R. Penuel, and A. Krause, “What is a “good” social network for policy implementation? The Flow of know-how for organizational change,” *Journal of Policy Analysis and Management*, vol. 34, no. 2, pp. 378–402, 2015.
- [10] L. Cai and X. Z. Yang, “Research on the application of big data in social public opinion monitoring and decision-making,” *Administrative Tribune*, vol. 11, p. 75, 2021.
- [11] H. Song, D. Peng, X. Huang, and J. Feng, “Research on Weibo hotspot finding based on self-adaptive incremental clustering,” *Journal of Shanghai Jiaotong University*, vol. 24, no. 3, pp. 364–371, 2019.
- [12] B. J. Ma, N. Zhang, and Q. Tan, “Analysis of influencing factors of public service effectiveness based on big data of interaction between government and people,” *Chinese Public Administration*, vol. 400, no. 10, pp. 111–117, 2018.
- [13] Y. Chun, Y. Kim, and H. Campbell, “Using Bayesian methods to control for spatial autocorrelation in environmental justice research: an illustration using toxics release inventory data for a Sunbelt county,” *Journal of Urban Affairs*, vol. 34, no. 4, pp. 419–439, 2012.
- [14] J. Li, Y. He, and Q. Xiong, “Research on network public opinion text mining based on big data technology,” *Journal of Intelligence*, vol. 13, no. 10, pp. 1–6, 2014.
- [15] T. Joachims, *Learning to Classify Text Using Support Vector Machines: Methods, Theory and algorithms*, Kluwer Academic Publishers, New York, NY, USA, 2002.
- [16] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent dirichlet allocation,” *Journal of Machine Learning Research*, vol. 3, no. 1, pp. 993–1022, 2003.
- [17] J. Weng, E. P. Lim, J. Jiang, and Q. He, “Twitterrank: finding topic-sensitive influential twitterers,” in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pp. 261–270, ACM, New York, NY, USA, February 2010.
- [18] C. Honey and S. C. Herring, “Beyond microblogging: conversation and collaboration via twitter system sciences,” in

Proceedings of the HICSS'09. 42nd Hawaii International Conference on, pp. 1–10, IEEE, Big Island, HI, USA, January 2009.

- [19] X. Yan, J. Guo, Y. Lan, and X. Cheng, “A biterm topic model for short texts,” in *Proceedings of the 22nd International Conference on World Wide Web*, pp. 1445–1456, ACM, New York, NY, USA, May 2013.
- [20] Q. Tang, *Short Text Clustering Based on BTM*, Anhui University, Hefei, China, 2014.
- [21] Y. Zhang, *Short Text Similarity Calculation Based on Feature Extension of BTM Topic Model*, Anhui University, Hefei, China, 2014.
- [22] A. Scrapy, *Fast and Powerful Scraping and Web Crawling Framework*, Scrapy. org. Np, Ballincollig, Ireland, 2016.

Research Article

An Autocontouring Method for the Kidneys Using an Adaptive Weighted Multikernel Support Vector Machines

Yi Gu ¹ and Bo Li^{1,2}

¹School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, Jiangsu, China

²Jiangsu Key Laboratory of Media Design and Software Technology, Jiangnan University, Wuxi 214122, Jiangsu, China

Correspondence should be addressed to Yi Gu; 8202101437@jiangnan.edu.cn

Received 19 April 2021; Accepted 5 July 2021; Published 14 July 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Yi Gu and Bo Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In radiotherapy, the location of the target area is very important. If the target area is small, the treatment is not complete, so the location of the target area is generally larger than the actual cancerous site. However, the damage of radiotherapy to normal cells is the same. In order to reduce the damage to the body as much as possible, we need to complete the most suitable target area. This paper uses an adaptive weighted multikernel support vector machine, which solves the parameter problem in the traditional multikernel support vector machine. The new AW-SVM can adjust the kernel weights adaptively. We completed our experiment on the abdominal MR dataset, using DSI as an evaluation indicator, and the experimental results showed its excellent classification performance. The minimum value of DSI in all results is 0.9654 (more than 0.7 is acceptable).

1. Introduction

In 2020, it is estimated that there will be 73750 new cases of kidney and renal pelvis cancer patients in the United States, of which 45520 are males, 28230 are females, and 14830 are expected to die, including 9860 males and 4970 females. In the five years from 2012 to 2016, the incidence of kidney and renal pelvis cancer was 16.6 per 100,000 people, 22.5 for men and 11.5 for women. In 2013–2017, kidney and renal pelvis cancer has a mortality rate of 3.7 per 100,000 people, 5.4 for men and 2.7 for women [1].

Adaptive radiation therapy (ART) is still the main method to solve cancer, especially for patients with advanced cancer [2]. Broadly speaking, any technology that adjusts the treatment process through feedback can be included in the scope of ART, such as image-guided radiation therapy (IGRT) [3], dose-guided radiation therapy (DGRT) [4], and structure-guided radiation therapy (SGRT) [5]. IGRT can be described as the primary stage of ART. It adds the concept of time factor on the basis of three-dimensional radiotherapy and fully considers the movement of anatomical tissue during the treatment process, and the displacement error of the divided

treatment time, such as breathing and peristaltic movement, daily pendulum position errors, and target area contraction, causes changes in the radiation dose distribution and imaging of the treatment plan. Before the patient undergoes treatment, various types of advanced imaging equipment are used to monitor the tumor and surrounding risk areas in real time and can adjust the treatment conditions according to the changes of the organs so that the irradiation field only follows the target area so that it can achieve precise treatment in the true sense. DGRT is proposed on the basis of IGRT. In addition to comparing image data, DGRT also compares the actual absorbed dose of the tumor and surrounding normal tissue during treatment with the dose in the treatment plan and adjusts the treatment plan in time. In general, adaptive ART is a self-responsive, self-correcting dynamic closed-loop system from diagnosis location, plan design, and treatment implementation to verification. Compared with computed tomography-guided ART (CT-ART), the main feature of MR-ART is that there is no radiation damage and no bone artifacts and it can perform multifaceted and multiparameter imaging, has a high degree of soft tissue resolution, can display vascular structures without the use

of contrast agents, and have multisequence biological function images [6, 7].

With the development of three-dimensional conformal radiotherapy and intensity-modulated radiotherapy in recent years, more and more researchers pay attention to the accuracy of tumor target definition to minimize the damage to organs at risk (OARs) [8]. The traditional circle drawing of the target area completely relies on manpower to complete it manually. Because manual contour drawing involves knowledge of multiple departments and engineering technology, it is extremely difficult for general doctors to draw. It is done by an experienced imaging surgeon or radiation oncologist. In most cases, it needs to be jointly discussed by multiple experts in the imaging department and the oncology department. However, there is also a problem. The manual contour is almost impossible to reproduce, and the target area contour depends entirely on the knowledge level of the experts and the image quality. There are also deviations between different experts. Manual contour drawing is also very time-consuming and labor-intensive. In the face of high-dimensional multimode images, as the images increase, the working pressure increases, which may cause contour deviations [9]. In addition, the accuracy of manual contours needs to be increased. It is unrealistic to use manual contours for online MR-ART, so it is necessary to develop an online automatic contour algorithm for MR-ART.

Broadly speaking, the current image segmentation algorithms mainly include segmentation methods based on edge detection, threshold, region growth, statistics, graph theory, information theory, fuzzy set theory, knowledge, and so on and segmentation methods based on convolutional neural networks.

Early image segmentation methods were mainly based on point, line, and edge detection and segmentation methods, using Robert operator [10], Canny operator [11], Sobel edge detection operator [12], and so on. Statistics-based segmentation methods include unsupervised k-means clustering, FCM clustering, and Markov random field, and the other is supervised support vector machine (SVM) [13], naive bayes (NB) [14], and random forest model [15]. Compared with the unsupervised segmentation algorithm, the supervised one obtains certain prior knowledge through training and performs better in image segmentation. But unfortunately, in the face of complicated and ambiguous organ boundaries, these algorithms have limited success in judging the boundaries.

With the continuous maturity and improvement of deep learning technology, the experimental results gradually exceed the previous machine learning, relying on the strong feature extraction ability of convolution neural network. The larger the amount of data given in the training process, the stronger the generalization performance of the network. In terms of image segmentation, FCN (Fully Convolutional Networks) [16] and U-net [17] are outstanding. FCN classifies images at the pixel level, thus solving the problem of image segmentation at the semantic level. U-net performs particularly well in medical image segmentation. U-net can make full use of the low-level and high-level information of medical images to provide a basis

for physical category recognition and precise segmentation and positioning.

The second chapter will introduce the dataset, data preprocessing, multiple kernel learning, and fixed-weight multikernel SVM to solve the problem of single-kernel inflexibility of SVM in image segmentation. Finally, we will introduce an adaptive weighted multikernel SVM to solve the weight problem of multikernel SVM. The third chapter will introduce the choice of kernel function, the evaluation index of the segmentation result, the comparison experiment setting, and the experiment process. The fourth chapter analyzes the experimental results, and the fifth chapter summarizes the full paper.

2. Related Work

2.1. Dataset. This experiment uses a dataset of eight patients with unresectable malignant tumors of the lower abdomen. We mark the eight patients as Sub1~8. Each patient's dataset consists of 16 pictures, and the first 15 pictures are in the process of treatment. The last one is the MRI image of the treatment day, the resolution is 370×370 , the axial pixel spacing is $1.5 \text{ mm} \times 1.5 \text{ mm}$, and the plane spacing is 3 mm. The total dataset consists of a total of 126 images of eight patients. The kidney and skin contours in the dataset are all drawn by professional radiation oncologists as our automatic contour assessment.

In terms of feature extraction, this article only uses the most basic voxel features and three-dimensional space coordinates. The local texture features (voxels) corresponding to each sequence are extracted from the four sequences shown in Figure 1. In terms of three-dimensional coordinates, since the coordinates of each sequence are the same, the voxel coordinates of the IP sequence are selected in this experiment.

2.2. Data Preprocessing. In order to further improve the separability of the data, consider using a filtering algorithm to process the data. The filtering algorithm filters the image information, removes the pixels we do not want, and enhances the information we need. Because the lower abdomen organs/body fluids of the human body are also constantly in motion, it is considered to use Kalman filter [18] to increase the dimension of the data. After experimental verification and analysis, it is completely appropriate to use Kalman filter to process this MR dataset.

Kalman filter is often used in uncertain systems (robots, real-time systems, etc.). It is relatively rare in classification algorithms. Kalman filter can infer the state of the next step based on the state of the previous step. In our MR data, the voxels are all adjacent, which is associated with this feature of Kalman filtering. Similar to the idea of KNN algorithm, adjacent voxels have the same label. Through experimental comparison, after using Kalman filter for feature extraction, the separability of the data has been greatly improved. The experimental results can be seen in Table 1.

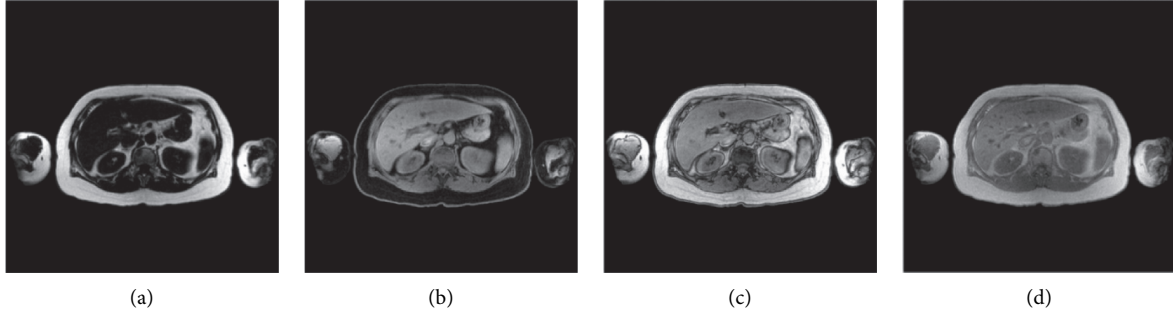


FIGURE 1: (a) Dixon-FAT sequence, (b) Dixon-WATER sequence, (c) Dixon-OP sequence, and (d) Dixon-IP sequence. All four sequences are reconstructed from Dixon sequence.

TABLE 1: Performance comparison of classifiers using Kalman filter.

<i>Part A: without a Kalman filter</i>				
Datasets	AW-SVM	SVM	S4VM	MKL
Sub1	0.9621 ± 0.0148	0.9643 ± 0.0014	0.9348 ± 0.0050	0.9652 ± 0.0139
Sub2	0.9811 ± 0.0102	0.9770 ± 0.0032	0.9524 ± 0.0046	0.9768 ± 0.0091
Sub3	0.9731 ± 0.0093	0.9712 ± 0.0013	0.9380 ± 0.0103	0.9725 ± 0.0124
Sub4	0.9875 ± 0.0115	0.9845 ± 0.0011	0.9525 ± 0.009	0.9794 ± 0.0121
Sub5	0.9848 ± 0.0122	0.9812 ± 0.0037	0.9553 ± 0.0033	0.9810 ± 0.0116
Sub6	0.9756 ± 0.0126	0.9746 ± 0.0021	0.9655 ± 0.0069	0.9727 ± 0.0112
Sub7	0.9818 ± 0.0116	0.9732 ± 0.0025	0.9610 ± 0.0026	0.9754 ± 0.0123
Sub8	0.9895 ± 0.088	0.9875 ± 0.0038	0.9507 ± 0.0113	0.9862 ± 0.0082
Datasets	RSM	TSVM	NB	RBF
Sub1	0.5549 ± 0.0036	0.7385 ± 0.0093	0.8646 ± 0.0070	0.6321 ± 0.0026
Sub2	0.5049 ± 0.0067	0.7886 ± 0.0028	0.8848 ± 0.0085	0.5894 ± 0.0037
Sub3	0.7060 ± 0.0082	0.7425 ± 0.0008	0.8667 ± 0.0071	0.5723 ± 0.004
Sub4	0.6415 ± 0.0058	0.7625 ± 0.0132	0.8949 ± 0.0039	0.6415 ± 0.0033
Sub5	0.6235 ± 0.0064	0.7625 ± 0.0116	0.9012 ± 0.0077	0.6149 ± 0.0037
Sub6	0.5936 ± 0.0081	0.8043 ± 0.0142	0.8980 ± 0.008	0.6512 ± 0.004
Sub7	0.5395 ± 0.0074	0.7389 ± 0.0092	0.8841 ± 0.0059	0.6035 ± 0.0034
Sub8	0.6043 ± 0.0098	0.7410 ± 0.0086	0.8468 ± 0.0073	0.9817 ± 0.0042
<i>Part B: with a Kalman filter</i>				
Datasets	AW-SVM	SVM	S4VM	MKL
Sub1	0.997 ± 0.0025	0.9966 ± 0.0015	0.9863 ± 0.0041	0.996 ± 0.0023
Sub2	0.9999 ± 0.0001	1	0.9888 ± 0.0048	0.9998 ± 0.0001
Sub3	0.9998 ± 0.0001	0.9995 ± 0.0003	0.9953 ± 0.0013	0.9995 ± 0.0004
Sub4	1	0.9999 ± 0.0001	0.9837 ± 0.0013	0.9999 ± 0.0001
Sub5	1	0.9999 ± 0.0001	0.9918 ± 0.021	1
Sub6	0.9999 ± 0.0001	1	0.9923 ± 0.0021	0.9997 ± 0.0002
Sub7	1	0.9999 ± 0.0001	0.9962 ± 0.0023	0.9994 ± 0.0003
Sub8	1	1	0.9835 ± 0.0026	1
Datasets	RSM	TSVM	NB	RBF
Sub1	0.8105 ± 0.0080	0.9646 ± 0.0075	0.9953 ± 0.0017	0.9817 ± 0.0034
Sub2	0.8641 ± 0.0031	0.9621 ± 0.0083	0.9267 ± 0.0084	0.9836 ± 0.0012
Sub3	0.9078 ± 0.0046	0.9784 ± 0.0071	0.9784 ± 0.0027	0.9951 ± 0.0034
Sub4	0.9033 ± 0.0069	0.9697 ± 0.0059	0.9172 ± 0.0077	0.9980 ± 0.0016
Sub5	0.8737 ± 0.074	0.9790 ± 0.0021	0.9556 ± 0.0112	0.9949 ± 0.0017
Sub6	0.8565 ± 0.0061	0.9641 ± 0.0043	0.9289 ± 0.0075	0.9823 ± 0.0025
Sub7	0.9159 ± 0.0034	0.9759 ± 0.0043	0.9589 ± 0.0045	0.9971 ± 0.0026
Sub8	0.8226 ± 0.0026	0.9135 ± 0.0095	0.8963 ± 0.0073	0.9941 ± 0.0032

2.3. *Multiple Kernel Learning.* We all know that SVM is a supervised learning method of machine learning. It is often used to solve classification problems. Its basic principle is to find a hyperplane in the feature space and separate positive and negative samples with the highest accuracy. However,

the previous SVMs are all single-kernel, which is based on a single feature space. In the actual classification task, we need our experience to select the appropriate kernel function (Gaussian kernel function, polynomial kernel function, etc.) and specify different parameters. This is not only

inconvenient, but also when our dataset features are heterogeneous, the effect is not so good.

The multiple kernel learning (MKL) [19] model is born because of this application. The multikernel model is more flexible than the single-kernel model and can give full play to the mapping capabilities of different kernel functions in a combined space composed of multiple feature spaces.

2.4. Multikernel SVM. According to the Mercer theorem and its properties [20], if K_1 and K_2 are kernels on $X \times X$ and X belongs to R , then the following are also kernel functions:

$$K(x, z) = K_1(x, z) + K_2(x, z), \quad (1)$$

$$K(x, z) = \alpha K_1(x, z). \quad (2)$$

If there is more than one kernel on $X \times X$, but multiple kernels coexist, then according to (1) and (2), we know that the following is also a kernel function:

$$K(x, z) = \sum_{m=1}^M d_m K_m(x, z), \quad d_m \geq 0, \quad (3)$$

$$\sum_{m=1}^M d_m = 1.$$

Equation (3) is the most common form of multiple kernel function combination, which is a weighted combination of single-kernel functions, is the basis kernel function, is the number of basis kernel functions, and is the weight coefficient corresponding to the first basis kernel function.

In general, the original problem of MKL can be described as follows:

$$\begin{cases} \min_{w_m, b, \xi, d} & \frac{1}{2} \sum_{m=1}^M \frac{1}{d_m} \|w_m\|_{H_m}^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} & y_i \left(\sum_{m=1}^M w_m \cdot \varphi(x_i) + b \right) \geq 1 - \xi_i, \quad \xi_i \geq 0 \\ & \sum_{m=1}^M d_m = 1, \quad d_m \geq 0, \end{cases} \quad (4)$$

where d_m is the kernel weight, w_m is the normal of the separating hyperplane corresponding to the M -th kernel, b is the deviation term, ξ is the slack variable, and C is the misclassification penalty coefficient. The above formula can be transformed into an optimization problem with d_m as a variable:

$$\begin{cases} \min_d & J(d) \\ \text{s.t.} & \sum_{m=1}^M d_m = 1, \quad d_m \geq 0, \end{cases}$$

$$J(d) = \begin{cases} \min_{w_m, b, \xi} & \frac{1}{2} \sum_{m=1}^M \frac{1}{d_m} \|w_m\|_{H_m}^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} & y_i \left(\sum_{m=1}^M w_m \cdot \varphi(x_i) + b \right) \geq 1 - \xi_i, \quad (i = 1, 2, \dots, n) \\ & \xi_i \geq 0. \end{cases} \quad (5)$$

The above objective function $J(d)$ is a standard SVM problem, which is a convex optimization problem that includes d_m . We can introduce a generalized Lagrangian

function to convert the original problem $J(d)$ into a dual problem. The dual problem is a minimax problem; then,

$$L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2} \sum_{m=1}^M \frac{1}{d_m} \|w_m\|_{H_m}^2 + C \sum_{i=1}^n \xi_i + \sum_{i=1}^n \alpha_i \left(1 - \xi_i - y_i \left(\sum_{m=1}^M w_m \cdot \varphi(x_i) + b \right) \right) - \sum_{i=1}^n \nu_i \xi_i, \quad (6)$$

where α and ν is the Lagrangian multiplier, and the Langrangian function $L(w_m, b, \xi, \alpha, \nu)$ takes the partial derivative of and makes it equal to zero, we can get

$$\begin{aligned}\nabla_{w_m} L(w_m, b, \xi, \alpha, \nu) &= \frac{1}{d_m} w_m - \sum_{i=1}^n \alpha_i y_i \varphi_m(x_i) = 0, \\ \nabla_b L(w_m, b, \xi, \alpha, \nu) &= -\sum_{i=1}^n \alpha_i y_i = 0, \\ \nabla_{\xi} L(w_m, b, \xi, \alpha, \nu) &= C - \alpha_i - \nu_i = 0.\end{aligned}\quad (7)$$

Through (8), we can get

$$\begin{aligned}w_m &= d_m \sum_{i=1}^n \alpha_i y_i \varphi_m(x_i), \\ \sum_{i=1}^n \alpha_i y_i &= 0, \\ C - \alpha_i + \nu &= 0.\end{aligned}\quad (8)$$

Substituting (9) into the Langrangian function (7), we can get

$$\begin{cases} \min_{w_m, b, \xi} L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \sum_{m=1}^M d_m K_m(x_i, x_j) + \sum_{i=1}^n \alpha_i \\ \text{s.t.} \quad \sum_{i=1}^n \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, n. \end{cases}\quad (9)$$

Find the maximum of $\min_{w_m, b, \xi}$ to α , and convert the maximum to the minimum to obtain the equivalent dual optimization problem:

$$\begin{cases} \min_{\alpha} L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \sum_{m=1}^M d_m K_m(x_i, x_j) - \sum_{i=1}^n \alpha_i \\ \text{s.t.} \quad \sum_{i=1}^n \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, n. \end{cases}\quad (10)$$

Formula (10) is a dual form of the standard SVM problem including the combined kernel $K(x_i, x_j)$, assuming that $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_n^*)^T$ is the optimal solution of the dual optimization problem, then for the expression

$$w_m^* = d_m \sum_{i=1}^n \alpha_i^* y_i^* \varphi_m(x_i), \quad (11)$$

$$b_m^* = y_j - \sum_{m=1}^M d_m \sum_{i=1}^n \alpha_i^* y_i K(x_i \cdot x_j). \quad (12)$$

Through (11) and (12), the final decision function is

$$f(x) = \text{sign} \left(\sum_{m=1}^M \left(d_m \sum_{i=1}^n \alpha_i^* y_i^* K(x \cdot x_i) + b_m^* \right) \right). \quad (13)$$

In summary, the description of the MKL learning algorithm can be found in Algorithm 1.

In this section, we introduced the multiple kernel learning. You can see that the kernel weights in MKL are fixed. In the next section, we will improve a new multiple kernel learning based on this point.

2.5. Adaptive Weighted Multikernel SVM. The kernel weight d_m in multikernel SVM is not adaptive, and it is more troublesome to use the gradient descent method to solve it. It takes multiple iterations to converge when it is close to the minimum. In this part we propose an adaptive weighted multikernel SVM model (AW-SVM).

In general, the original problem of AW-SVM is as follows:

$$\begin{cases} \min_{w_m, b, \xi, d} \frac{1}{2} \sum_{m=1}^M d_m^r \|w_m\|_{H_m}^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad y_i \left(\sum_{m=1}^M w_m \cdot \varphi(x_i) + b \right) \geq 1 - \xi_i, \quad \xi_i \geq 0 \\ \sum_{m=1}^M d_m = 1, \quad d_m \geq 0, \end{cases}\quad (14)$$

where d_m^r is the kernel weight, the index r is similar to the fuzzy index in FCM clustering, which is a relaxation of

Input: training set: $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ where $x_i \in \chi \in R^n, y_i \in \{-1, +1\}, i = 1, 2, \dots, N$
Output: classification decision function $f(x)$
Step 1: choose the appropriate kernel function $K(x, z)$, the appropriate parameter C , and the loss parameter ξ
Kernel weight matrix d_m initialization: $d_m^1 = 1/M$, where M is the number of kernels
Construct and solve optimization problems (10), and find the optimal solution $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_l^*)^T$
Step 2: choose $0 < \alpha^* < C$ positive component α from α^* , and calculate formula (11) and formula (12)
Step 3: construct decision function (13)

ALGORITHM 1: Learning algorithm for MKL.

the weight, and is used to realize the self-adaptation of the weight, which w_m is the normal of the separating hyperplane corresponding to the M -th kernel, b is the deviation term, ξ

is the slack variable, and C is the misclassification penalty coefficient. Formula (14) can be transformed into an optimization problem with d_m as a variable

$$\begin{cases} \min_d & J(d) \\ \text{s.t.} & \sum_{m=1}^M d_m = 1, \quad d_m \geq 0, \end{cases}$$

$$J(d) = \begin{cases} \min_{w_m, b, \xi} & \frac{1}{2} \sum_{m=1}^M d_m^r \|w_m\|_{H_m}^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} & \left(\sum_{m=1}^M w_m \cdot \varphi(x_i) + b \right) \geq 1 - \xi_i, \quad (i = 1, 2, \dots, n) \\ & \xi_i \geq 0. \end{cases} \quad (15)$$

The above objective function $J(d)$ is a standard SVM problem, which is a convex optimization problem involving d_m . Using the same method, we can get

$$L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2} \sum_{m=1}^M d_m^r \|w_m\|_{H_m}^2 + C \sum_{i=1}^n \xi_i + \sum_{i=1}^n \alpha_i \left(1 - \xi_i - y_i \left(\sum_{m=1}^M w_m \cdot \varphi(x_i) + b \right) \right) - \sum_{i=1}^n \nu_i \xi_i. \quad (16)$$

Use the same method in the previous section (7), (8), and (10) to get the final dual optimization problem:

$$\begin{cases} \min_{\alpha} & L(w_m, b, \xi, \alpha, \nu) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \sum_{m=1}^M \frac{1}{d_m^r} K_m(x_i, x_j) - \sum_{i=1}^n \alpha_i \\ \text{s.t.} & \sum_{i=1}^n \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, n. \end{cases} \quad (17)$$

Assuming that $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_l^*)^T$ is the optimal solution of the dual optimization problem, then for the expression

$$w_m^* = \frac{1}{d_m^r} \sum_{i=1}^n \alpha_i^* y_i^* \varphi_m(x_i), \quad (18)$$

$$b_m^* = y_j - \sum_{m=1}^M \frac{1}{d_m^r} \sum_{i=1}^n \alpha_i^* y_i^* K(x_i \cdot x_j). \quad (19)$$

$$J(d, \alpha, \nu) = \frac{1}{2} \sum_{m=1}^M d_m^r \|w_m\|_{H_m}^2 + C \sum_{i=1}^n \xi_i + \sum_{i=1}^n \lambda_i \left(1 - \xi_i - y_i \left(\sum_{m=1}^M w_m \cdot \varphi(x_i) + b \right) \right) - \sum_{i=1}^n \nu_i \xi_i + \lambda_1 \left(\sum_{m=1}^M d_m - 1 \right). \quad (20)$$

Taking the partial derivative of the parameter d_m, λ_1 and making it equal to zero, we can get

$$\frac{\partial J(d)}{\partial (d_m)} = \frac{r}{2} d_m^{r-1} \|w_m\|^2 + \lambda_1 = 0, \quad (21)$$

$$\frac{\partial J(d)}{\partial (\lambda_1)} = \sum_{m=1}^M d_m = 1. \quad (22)$$

From the simultaneous formulas (21) and (22), using a similar solution method of U_{ij}^m in FCM [21], the final solution d_m is equal to

$$d_m = \frac{\left(1/\|w_m\|^2 \right)^{(1/r-1)}}{\sum_{k=1}^M \left(1/\|w_k\|^2 \right)^{(1/r-1)}}. \quad (23)$$

In summary, the two classification problem of AW-SVM can be described as follows:

Step 1: initialization of kernel weight matrix d_m , $\sum_{m=1}^M d_m = 1$, M is the number of kernels.

After getting the parameter w_m , d_m can be solved by iteration.

To construct a Lagrangian function with d_m as a variable for the original problem, we get

Step 2: in each iteration, use the combined kernel $K = \sum_{m=1}^M 1/d_m^r K_m$ to calculate the final dual optimization problem (17).

Step 3: calculate the optimal solution of (17). Calculate the separation hyperplane (w_m, b) corresponding to each kernel by formulas (18) and (19).

Step 4: update d_m by formula (23).

Step 5: if the iteration termination condition is not met, return to Step 2 and repeat Steps 2–4. If the iteration condition is satisfied, the calculation is ended. Then, output the final decision function:

$$f(x) = \text{sign} \left(\sum_{m=1}^M \frac{1}{d_m^r} \left(\sum_{i=1}^n \alpha_i^* y_i^* K(x \cdot x_i) + b_m^* \right) \right). \quad (24)$$

Iterative stop condition is duality gap (DG) [22]. Karush–Kuhn–Tucker (KKT) condition, $\Delta d = d^{t+1} - d^t \leq \varepsilon_1$, reaches the maximum number of iterations.

The DG expression is

$$\max_m \sum_{i=1}^n \sum_{j=1}^n \alpha_i^* \alpha_j^* y_i y_j K_m(x_i \cdot x_j) - \sum_{i=1}^n \sum_{j=1}^n \alpha_i^* \alpha_j^* y_i y_j \sum_{m=1}^M \frac{1}{d_m^r} K_m(x_i, x_j) \leq \varepsilon. \quad (25)$$

Both ε and ε_1 are thresholds.

3. Experiment

After getting the artificial contour image of all the data drawn by the expert, we start to train the AW-SVM learning algorithm. The algorithm runs on a computer with Intel(R) Core(TM) i5-8500CPU 3.00 GHZ 12 GB memory and 64 bit Windows10 operating system. The algorithm iterates 10

times on average, and the training time is about 30 minutes. Given a trained classifier model, the average segmentation time of images on treatment days is about 3 minutes.

3.1. Kernel Function. In this experiment, we choose Gaussian kernel function (26), Laplacian kernel function (27), and logarithmic kernel function (28). The formulas of the three are as follows:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right), \quad (26)$$

$$K(x, y) = \exp\left(-\frac{\|x - y\|}{\sigma}\right), \quad (27)$$

$$K(x, y) = -\log(1 + \|x - y\|^d). \quad (28)$$

The Gaussian kernel is a classic robust radial basis kernel. It has good anti-interference ability against the noise in the data, which has been widely used, but the performance of the Gaussian kernel function is very sensitive to the parameters. The Laplacian kernel function is a variant of the Gaussian kernel function. The main change is to adjust the 2-norm to 1-norm, which reduces the sensitivity to parameters. Logarithmic kernel is often used in image segmentation.

3.2. Evaluation Index. In addition to using classification accuracy to evaluate the accuracy of the algorithm, this article also uses Dice Loss [23], which is very common in segmentation algorithms, to evaluate our segmentation accuracy, because segmentation usually classifies each pixel:

$$\text{Dice Loss} = 1 - \text{Dice Coefficient}. \quad (29)$$

Dice Coefficient, named after Lee Raymond Dice, is a set similarity measure function, which is usually used to calculate the similarity between two samples (value $\in [0, 1]$). A value of 1 indicates complete coincidence. The formula is as follows:

$$\text{Dice Coefficient} = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (30)$$

where $|X|$ and $|Y|$, respectively, represent the number of elements in the set, as shown in Figure 2. In the segmentation task, the two, respectively, represent the real result and the predicted result.

Combining (29) and (29), Dice Loss expression is as follows:

$$\text{Dice Loss} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}. \quad (31)$$

3.3. Comparative Test. In order to compare the performance of the AW-SVM algorithm on this dataset, this paper arranges other seven algorithms as an experimental comparison. The comparison algorithm is shown in Table 2.

The experimental training set consists of daily MRI data before treatment, and the treatment day dataset is used as the test set. The organs and skin except the kidney in the dataset are labeled as background label 1, and the kidney as the circle target, and the label is -1 .

3.4. Experiment Process. For each patient (Sub1~8), complete the preliminary data processing (local texture feature extraction, three-dimensional coordinate extraction, Kalman filter dimension enhancement, etc.). The MRI image

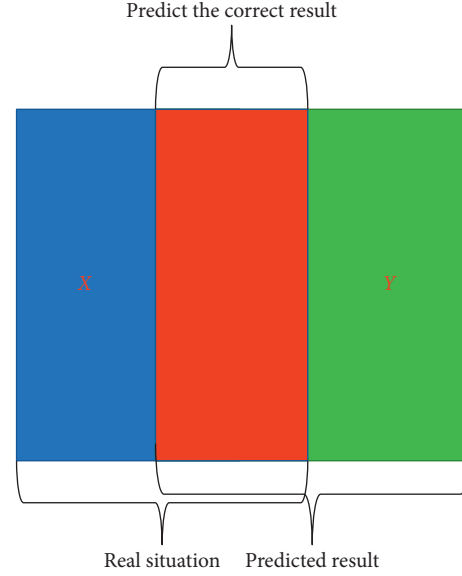


FIGURE 2: The description of the Dice Coefficient.

TABLE 2: Comparison algorithm used.

Our algorithm	Comparison algorithm
Adaptive weighted multikernel SVM (AW-SVM)	Support vector machine (SVM)
	Multiple kernel learning (MKL)
	Transductive support vector machine (TSVM) [24]
	Semisupervised SVM (S4VM) [25]
	Random subspace method (RSM) [26]
	Naive bayes (NB)
	Radial basis function (RBF) [27]

during the treatment will be used as training data, denoted as I_t ; the MRI image of the treatment day is used as the test data, denoted as I_s . The experiment is carried out as shown in Figure 3:

- (1) Randomly select the $L = 4000$ group of examples in the labeled I_t (label: $\{-1, 1\}$); that is, select 4000 voxels from the kidney and background in the MRI image during the treatment process to form the training set G .
- (2) Using the AW-SVM method on the training set G , through parameter tuning, we can get the best classifier C_r for this patient data.
- (3) In the treatment day MRI dataset I_t , select a complete treatment day MRI image as the test input (including 36,000+ voxels), the label corresponding to each voxel is obtained through the best classifier C_r trained in Step 2, and the result is recorded as U .
- (4) Select all MRI images of the treatment day on I_t as the test input, and obtain the labels corresponding to all voxels through the KNN algorithm (label in U is used as the training label).

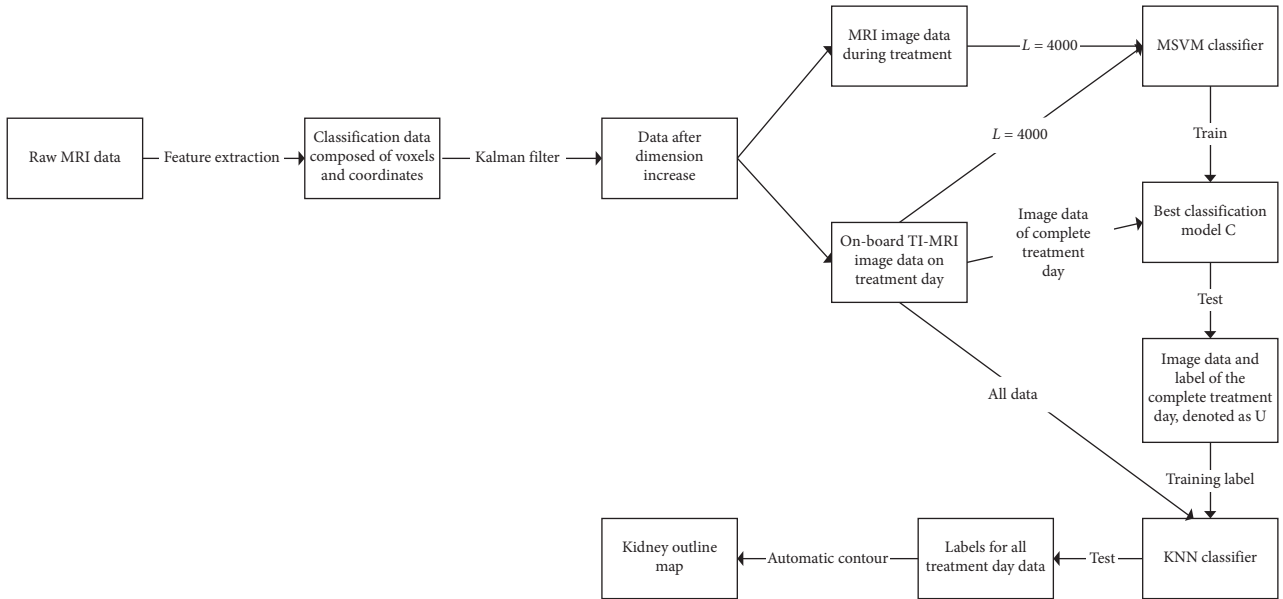


FIGURE 3: Experiment process.

- (5) Obtain the outline of the kidney through the obtained label.

3.5. Results and Discussion. In order to verify that the data processed by Kalman filter performs better in classification performance, we set up the first set of comparative experiments, we have prepared two sets of patient data; the one that is not processed by Kalman filtering is recorded as the first group, and the one that is processed by Kalman filtering is recorded as the second group. The eight algorithms have been experimented on these two sets of data, and the experimental results are shown in Table 1, from which we can get the following information:

- (1) It can be clearly seen that compared with the first set of results, each classifier performs better on the second set of data, and the classification accuracy has been improved, indicating that the use of Kalman filtering can improve the classification accuracy and enhance interpretable.
- (2) In each set of results, basically the classification results of AW-SVM are the best. Classification accuracy is much higher than TSVM, S4VM, NB, RBF, and RSM, and the accuracy of MKL and SVM is close to AW-SVM, but for this sample, AW-SVM's performance is still the best. In general, the experimental results show that our algorithm is effective and excellent for classification on the MRI data of the lower abdomen.

In order to compare whether there are significant differences between AW-SVM and the other seven algorithms in statistics, Friedman test [28, 29] is used to analyze the classification accuracy of the eight algorithms. In the Friedman test, each individual has the same sample size, and each individual has a corresponding relationship with the

corresponding individual in other samples, so the Friedman test can make full use of the information in the original data. If $p \leq 0.025$, the null hypothesis was rejected. At the same time, in order to further verify that the method to obtain the best Friedman ranking and other methods are statistically significant differences, we also conducted Holm post hoc test [28, 29].

As shown in Tables 3 and 4, the Friedman test shows that the classification performance of the AW-SVM method in this dataset is greater than that of the other seven algorithms. Holm post hoc test shows that the performance of AW-SVM algorithm on this dataset is significantly better than RSM, NB, S4VM, TSVM, and RBF, and for S4VM and MKL, it is not obvious. In summary, the AW-SVM algorithm can be used for lower abdomen MRI image segmentation, and its classification performance is better than the comparison of these seven classification algorithms. The experimental results show that its accuracy rate is close to 1 on each dataset, and the subsequent DSI values also show its excellent classification performance.

In the field of medical segmentation, Dice similarity index (DSI) value greater than 0.7 is an acceptable value, which means that the predicted contour and the real contour have a good overlap [30]. The results of DSI value are shown in Table 5. The performance of AW-SVM, SVM, and MKL models is as good as accuracy; DSI value reached 1 in multiple patients, which indicates that the automatic contour and manual contour are completely coincident. The lowest DSI values of the three are generated on Sub1, respectively, 0.9654 ± 0.0121 , 0.9423 ± 0.0127 , and 0.9324 ± 0.0162 , which are also acceptable values, but it also shows that the MRI image of Sub1 is the most complicated (there may be overlaps, missing, shadows, etc.) [31]. The performance of several other algorithms is uneven, but the DSI value of the RSM algorithm is around 0.2, which is a completely

TABLE 3: Result of Holm's post hoc test ($\alpha = 0.05$).

Holm's post hoc test for AW-SVM				
Algorithms	$z = (R_o - R_i)/SE$	P	holm = α/i	Hypothesis
RSM	5.358259	0	0.007143	Rejected
NB	4.184545	0.000029	0.008333	Rejected
TSVM	3.878359	0.000105	0.01	Rejected
S4VM	2.602583	0.009252	0.0125	Rejected
RBF	2.602583	0.009252	0.016667	Rejected
MKL	0.918559	0.358326	0.025	Not rejected
SVM	0.459279	0.646034	0.05	Not rejected

TABLE 4: Result of Friedman test ($\alpha = 0.05$).

Friedman test for AW-SVM			
Algorithm	Friedman rank	P value	Hypothesis
RSM	8		
NB	6.5625		
TSVM	6.1875		
S4VM	4.625	0	Rejected
RBF	4.625		
MKL	2.5625		
SVM	2		
AW-SVM	1.4375		

TABLE 5: Average dice similarity index (DSI) for the kidneys using the classification algorithm.

Part A: enter as part of a complete treatment day image				
Datasets	AW-SVM	SVM	S4VM	MKL
Sub1	0.9654 \pm 0.0121	0.9423 \pm 0.0127	0.7176 \pm 0.0211	0.9324 \pm 0.0162
Sub2	1	1	0.8355 \pm 0.0156	1
Sub3	0.9882 \pm 0.0094	0.9749 \pm 0.0114	0.8431 \pm 0.0257	0.9782 \pm 0.0076
Sub4	1	1	0.9100 \pm 0.0336	0.9995 \pm 0.0002
Sub5	1	1	0.8800 \pm 0.0546	1
Sub6	1	1	0.8436 \pm 0.0235	0.9988 \pm 0.0011
Sub7	1	1	0.8852 \pm 0.0303	1
Sub8	1	1	0.7181 \pm 0.0489	1
Datasets	RSM	TSVM	NB	RBF
Sub1	0.2157 \pm 0.0021	0.4267 \pm 0.0153	0.7510 \pm 0.0189	0.5976 \pm 0.0213
Sub2	0.3078 \pm 0.0012	0.6127 \pm 0.0483	0.8298 \pm 0.0220	0.6624 \pm 0.0176
Sub3	0.1724 \pm 0.0031	0.6925 \pm 0.0221	0.8528 \pm 0.0176	0.6768 \pm 0.0229
Sub4	0.2226 \pm 0.0012	0.8986 \pm 0.0318	0.9220 \pm 0.0304	0.7633 \pm 0.0318
Sub5	0.1862 \pm 0.0016	0.8318 \pm 0.0430	0.9111 \pm 0.019	0.7796 \pm 0.0174
Sub6	0.2543 \pm 0.0021	0.6811 \pm 0.0194	0.8363 \pm 0.0343	0.6717 \pm 0.0312
Sub7	0.2784 \pm 0.0011	0.5653 \pm 0.0236	0.8968 \pm 0.0308	0.6602 \pm 0.0158
Sub8	0.1748 \pm 0.0016	0.7168 \pm 0.0297	0.7113 \pm 0.0352	0.8681 \pm 0.0153

unacceptable value. Indicating that RSM is not applicable on this dataset.

Figure 4 shows the original cross-sectional image of the lower abdomen generated by the OP sequence (the clearest OP sequence) in MRI, the manual contour drawn by a professional doctor, and the automatic contour generated based on the AW-SVM algorithm. The manual contour is used as the evaluation criterion.

The running time of the automatic contour method is also a key point for clinical application. The faster the treatment speed, the shorter the time the patient will be fixed on the treatment bed and the less pain they will suffer. We

select the three models with the best performance of DSI and calculate their prediction time; the prediction time refers to the time it takes to input the complete treatment day data and get the automatic contour after obtaining the corresponding best classification model. The prediction time is shown in Table 6, affected by programming; our AW-SVM algorithm is not superior in predicting time, but it is also a completely acceptable value, with an average time of 2.9263 seconds. The automatic contour algorithm based on AW-SVM greatly shortens the drawing of the contour of the target area on the treatment day and greatly reduces the pain of the patient.

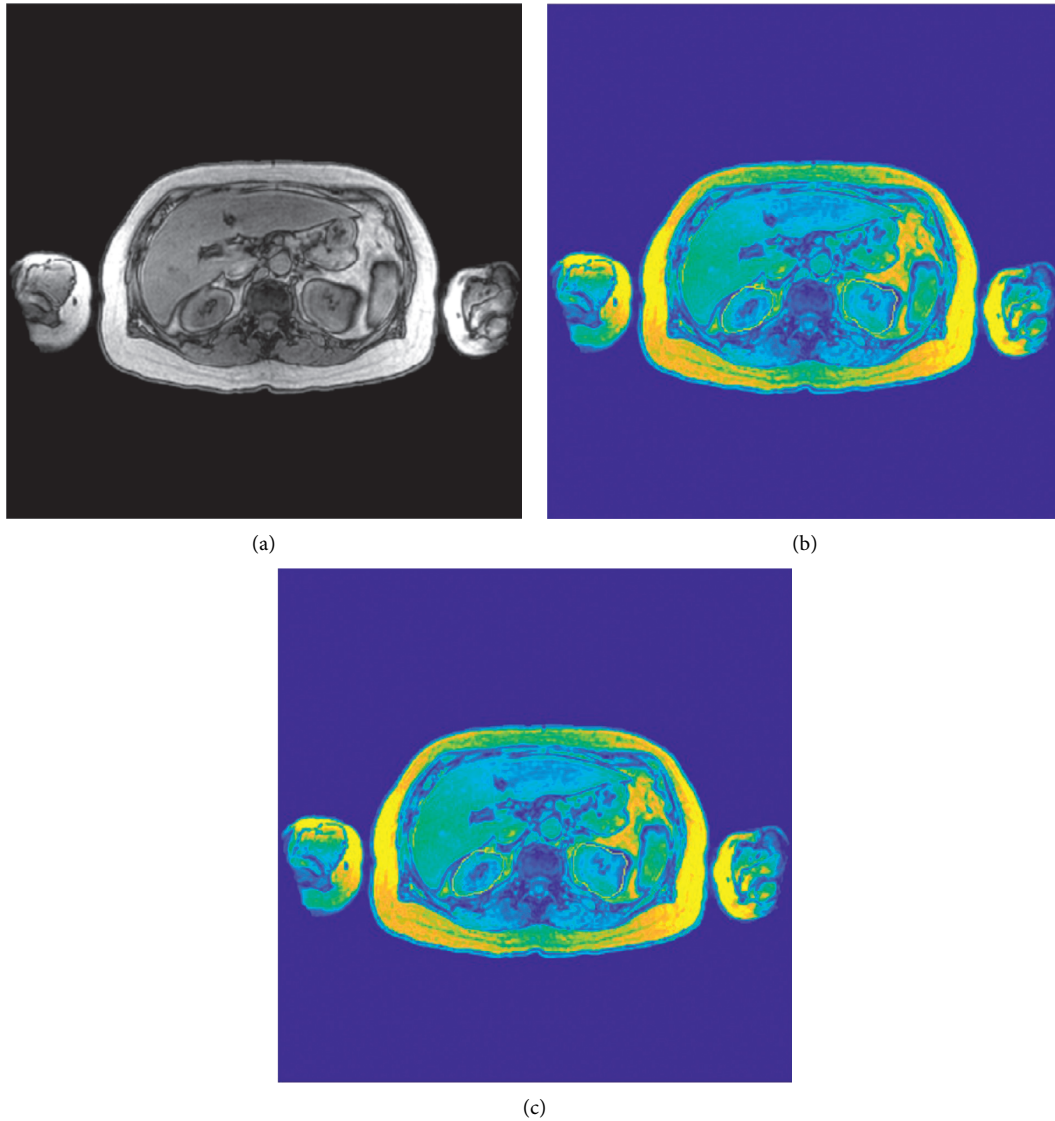


FIGURE 4: Abdominal MRI cross-sectional image showing the automatic contour for kidneys obtained by the AW-SVM method. (a) Original image, (b) manual contour, and (c) automatic contour. Red: manual contours used for ground truth; green: automatic contours of the kidneys.

TABLE 6: Average prediction time of the AW-SVM and MKL and SVM methods.

Datasets	AW-SVM	MKL	SVM
Sub1	2.53	1.18	0.23
Sub2	3.01	1.81	0.24
Sub3	2.87	1.56	0.22
Sub4	2.81	1.79	0.26
Sub5	2.73	1.27	0.25
Sub6	3.12	1.22	0.27
Sub7	3.02	1.31	0.25
Sub8	3.32	1.20	0.29

4. Conclusion

A mature automatic contour algorithm is the key to the clinical application of MR-ART. The experimental results

show that the adaptive weighted AW-SVM algorithm proposed in this paper shows excellent performance in abdominal MRI image segmentation, which is comparable to many excellent classification algorithms. The features

used in this article are only simple voxel and coordinate features. In the future, deep neural networks can be used to process data to extract deep features to further improve the classification performance of the AW-SVM algorithm. In the future, it will not only be drawn for a single kidney outline but also is hoped that all organs in the abdomen can be drawn.

Data Availability

The labeled datasets 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.

Authors' Contributions

Yi Gu and Bo Li contributed equally to this work.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61772241 and in part by the 2018 Six Talent Peaks Project of Jiangsu Province under Grant XYDXX-127.

References

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.
- [2] A. Brandmaier and S. C. Formenti, "The impact of radiation therapy on innate and adaptive tumor immunity," *Seminars in Radiation Oncology*, vol. 30, no. 2, pp. 139–144, 2020.
- [3] R. K. Funk, A. L. Stockham, and N. N. I. Laack, "Basics of radiation therapy," *Clinical Cardio-Oncology*, Elsevier, Amsterdam, Netherland, 2016.
- [4] J. Chen, O. Morin, M. Aubin, M. K. Bucci, C. F. Chuang, and J. Pouliot, "Dose-guided radiation therapy with megavoltage cone-beam CT," *The British Journal of Radiology*, vol. 79, no. 1, pp. 87–98, 2006.
- [5] S. Kuznetsova, P. Grendarova, S. Roy, R. Sinha, K. Thind, and N. Ploquin, "Structure guided deformable image registration for treatment planning CT and post stereotactic body radiation therapy (SBRT) Primovist (Gd-EOB-DTPA) enhanced MRI," *Journal of Applied Clinical Medical Physics*, vol. 20, no. 12, pp. 109–118, 2019.
- [6] B. W. Raaymakers, J. J. W. Lagendijk, J. Overweg et al., "Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept," *Physics in Medicine and Biology*, vol. 54, no. 12, pp. N229–N237, 2009.
- [7] S. Devic, "MRI simulation for radiotherapy treatment planning," *Medical Physics*, vol. 39, no. 11, pp. 6701–6711, 2012.
- [8] S. Acharya, B. W. Fischer-Valuck, R. Kashani et al., "Online magnetic resonance image guided adaptive radiation therapy: first clinical applications," *International Journal of Radiation Oncology*Biophysics*Physics*, vol. 94, no. 2, pp. 394–403, 2016.
- [9] J. Yang, H. Veeraraghavan, W. Elmpt, A. Dekker, M. Gooding, and G. Sharp, "CT images with expert manual contours of thoracic cancer for benchmarking auto-segmentation accuracy," *Medical Physics*, vol. 47, no. 7, pp. 3250–3255, 2020.
- [10] L. G. Roberts, *Machine Perception of Three-Dimensional Solids*, Massachusetts Institute of Technology, Cambridge, England, 1965.
- [11] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.
- [12] I. Sobel, "Camera models and machine perception," dissertation, stanford university, 1970.
- [13] C. J. C. Burges, "A tutorial on Support Vector Machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, pp. 121–167, 1998.
- [14] K. Adi, C. E. Widodo, A. P. Widodo, R. Gernowo, A. Pamungkas, and R. A. Syifa, "Naïve Bayes algorithm for lung cancer diagnosis using image processing techniques," *Advanced Science Letters*, vol. 23, no. 3, pp. 2296–2298, 2017.
- [15] S. L. Salzberg, *Book Review: C4.5: Programs for Machine Learning*, J. Ross Quinlan, Ed., Springer, New York, NY, USA, 1993.
- [16] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2015.
- [17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 9351, pp. 234–241, Springer International Publishing, Munich, Germany, October 2015.
- [18] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [19] M. Gönen and E. Alpaydın, "Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, pp. 2211–2268, 2011.
- [20] J. Mercer, "XVI. Functions of positive and negative type, and their connection the theory of integral equations," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 209, no. 441–458, pp. 415–446, 1909.
- [21] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, 2002.
- [22] A. Rakotomamonjy, F. R. Bach, S. Canu, and Y. Grandvalet, "Simple MKL," *Journal of Machine Learning Research*, vol. 9, no. 3, pp. 2491–2521, 2008.
- [23] F. Milletari, N. Navab, and S. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the Fourth International Conference on 3D Vision*, pp. 565–571, 3DV), Stanford, CA, USA, October 2016.
- [24] O. Chapelle, V. Sindhwani, and S. S. Keerthi, "Optimization techniques for semi-supervised support vector machines," *Journal of Machine Learning Research*, vol. 9, no. 6, pp. 203–223, 2008.
- [25] Z. Xu, Y. Tian, Z. Li, and Y. Li, "SEMG multi-class classification based on S4VM algorithm," *Lecture Notes in Electrical Engineering*, vol. 21, no. 4, pp. 495–503, 2013.
- [26] N. García-Pedrajas and D. Ortiz-Boyer, "Boosting random subspace method," *Neural Networks*, vol. 21, no. 9, pp. 1344–1362, 2008.

- [27] J. Park and I. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 2014.
- [28] J. Demiar and D. Schuurmans, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, no. 1, pp. 1–30, 2006.
- [29] S. Gar'ci and F. Herrera, "An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons," *Journal of Machine Learning Research*, vol. 9, pp. 2677–2694, 2008.
- [30] K. H. Zou, S. K. Warfield, A. Bharatha et al., "Statistical validation of image segmentation quality based on a spatial overlap index1," *Academic Radiology*, vol. 11, no. 2, pp. 178–189, 2004.
- [31] L. Fan, P. Qian, K.-H. Su et al., "Abdominal, multi-organ, auto-contouring method for online adaptive magnetic resonance guided radiotherapy: an intelligent, multi-level fusion approach," *Artificial Intelligence in Medicine*, vol. 90, pp. 34–41, 2018.

Research Article

Stock Price Forecast Based on CNN-BiLSTM-ECA Model

Yu Chen ¹, Ruixin Fang ¹, Ting Liang ¹, Zongyu Sha ¹, Shicheng Li ¹, Yugen Yi ¹,
Wei Zhou ², and Huilin Song ³

¹School of Software, Jiangxi Normal University, Nanchang 330022, China

²School of Computer, Shenyang Aerospace University, Shenyang 110136, China

³School of International Economics and Trade, Jiangxi University of Finance and Economics, Nanchang 330022, China

Correspondence should be addressed to Yugen Yi; yiyg510@jxnu.edu.cn and Wei Zhou; zhouweineu@outlook.com

Received 4 May 2021; Revised 3 June 2021; Accepted 28 June 2021; Published 9 July 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Yu Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Financial data as a kind of multimedia data contains rich information, which has been widely used for data analysis task. However, how to predict the stock price is still a hot research problem for investors and researchers in financial field. Forecasting stock prices becomes an extremely challenging task due to high noise, nonlinearity, and volatility of the stock price time series data. In order to provide better prediction results of stock price, a new stock price prediction model named as CNN-BiLSTM-ECA is proposed, which combines Convolutional Neural Network (CNN), Bidirectional Long Short-term Memory (BiLSTM) network, and Attention Mechanism (AM). More specifically, CNN is utilized to extract the deep features of stock data for reducing the influence of high noise and nonlinearity. Then, BiLSTM network is employed to predict the stock price based on the extracted deep features. Meanwhile, a novel Efficient Channel Attention (ECA) module is introduced into the network model to further improve the sensitivity of the network to the important features and key information. Finally, extensive experiments are conducted on the three stock datasets such as Shanghai Composite Index, China Unicom, and CSI 300. Compared with the existing methods, the experimental results verify the effectiveness and feasibility of the proposed CNN-BiLSTM-ECA network model, which can provide an important reference for investors to make decisions.

1. Introduction

With the unprecedented development of the network, multimedia data such as text, image, video, financial data from mobile phones, social networking sites, news, and financial websites are growing at a rapid pace and also affecting our real daily life. In the era of big data, how to make full use of these data providing the relevant and valuable information for us becomes a great significance task [1, 2]. For example, investors can employ the financial data to predict the future price trend of financial assets to reduce the decision-making risk [3, 4]. However, investors can hardly acquire the useful information of the budget allocation timely. In order to make right investment decisions for investors, some technical or quantitative methods are

necessary and important to use to predict the fluctuation of asset prices [5].

There are many methods that have been proposed to predict the financial stock data, which achieve excellent performance [5, 6]. For instance, traditional methods based on econometric statistical model aim to find the best estimation for time series prediction, such as Auto Regressive (AR) model, Moving Average (MA) model, Auto Regressive and Moving Average (ARMA) model, and Autoregressive Integrated Moving Average (ARIMA) model [7]. Although the abovementioned approaches can describe and evaluate the relationship between variables through statistical inference, there are still some limitations. On one hand, since these methods are based on the assumption of linear relationship of model structure, they can hardly capture the

nonlinear variation of the stock price [8, 9]. On the other hand, these approaches assume that the data have constant variance, while the financial time series have high-noisy, time-varying, dynamic properties, and so on [10].

In order to solve the aforementioned shortcomings, many machine learning technologies have been applied to simulate the nonlinear relationship in financial time series. Among them, Artificial Neural Network (ANN) owing to its excellent ability in nonlinear mapping and generalization has been widely used for dealing with the financial time series [11, 12]. Different from the econometric statistical models, the strict model structure and additional series of assumptions are not required for ANN model. For example, Hajizadeh et al. [13] put forward a hybrid model, which combined ANN with Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model to predict the volatility of S&P500. The experimental results show that the test error of the hybrid model is lower than that of any single econometric model. Rather et al. [14] realized stock return prediction via taking ARMA and Exponential Smoothing (SE) linear models with Recurrent Neural Network (RNN) [15] into consideration. The experimental results indicate that RNN can further improve the prediction performance.

In recent years, Long Short-Term Memory (LSTM) network has been proposed which overcomes the vanishing gradient problem in RNN and uses storage cells and gates to learn long-term correlation in time series data [16]. LSTM has been widely used for forecasting time series due to its advantages [17–19]. Nelson et al. [20] proposed a stock price prediction method based on LSTM. From the experimental results [20], we can find that the LSTM model is more accurate than other machine learning models, such as Support Vector Machine (SVM) [21], Genetic Algorithm (GA) [22], and BP Neural Network [23, 24]. Graves and Schmidhuber [25] proposed a Bidirectional LSTM (BiLSTM) network. BiLSTM contains a forward LSTM and a reverse LSTM, in which the forward LSTM takes advantage of past information and the reverse LSTM takes advantage of future information. Then, many researchers applied BiLSTM to solve the problem of time series prediction [26–28]. Because BiLSTM network can use the information of both the past and the future, the final prediction is more accurate than unidirectional LSTM. In addition, multitask RNNs methods also have been proposed for forecasting time series such as EEG-based motion intention recognition and dynamic illness severity prediction [29, 30].

Attention Mechanism (AM) is derived from human vision [31]. By scanning the target area quickly, the human eyes can pay more attention on the key areas to obtain the useful information and suppress other useless information. Thus, we can use the limited resources to quickly screen out the valuable information based on the target from the massive information [31]. At present, attention mechanism has been well applied in computer vision question answering [32, 33], natural language processing [34, 35], and speech recognition [36, 37]. In addition, some researchers have also successfully applied the attention mechanism to the related research of time series [38–41].

Inspired by the successful applications of deep learning and attention mechanism on stock data analysis [42–44], this paper proposes a time series prediction model named as CNN-BiLSTM-ECA, which integrates Convolutional Neural Networks (CNN) [10, 12] and BiLSTM to predict the closing price of stock data. First, our model adopts CNN to extract the deep features. Second, the feature vector is constructed in time series regarded as the input of the BiLSTM network. Third, in order to further improve the prediction performance of BiLSTM network, a novel attention mechanism termed as Effective Channel Attention (ECA) [45] module is introduced. Compared with other attention mechanisms, ECA is lighter and less complex, which can greatly improve network performance. Finally, we compare the proposed CNN-BiLSTM-ECA model with the LSTM, BiLSTM, and CNN-BiLSTM models on three stocks' data including Shanghai Composite Index, China Unicom, and CSI 300 to verify its effectiveness.

The outline of this paper is as follows: Section 2 reviews the related works and backgrounds are introduced in Section 3. Section 4 gives the proposed network model structure in detail. Section 5 shows extensive experiments to prove the effectiveness of the proposed approach. Section 6 presents some conclusions.

2. Related Works

Recently, a large number of stock prediction methods have been proposed. This paper mainly reviews two kinds of methods based on machine learning and deep learning.

2.1. Machine Learning Methods. As machine learning techniques are becoming more and more popular, Machine Learning (ML) methods based on financial time series forecasting have been studied extensively. Specifically, Nayak et al. [46] proposed a hybrid model based on SVM and K-Nearest Neighbor (KNN) to predict the Indian Stock Market Index. Combining weight SVM and KNN, Chen and Hao [47] proposed a new model for predicting the trend of the Chinese stock market. In order to find the optimal solution of neural network, Chiang et al. [48] developed a novel model to predict stock trend by introducing the Particle Swarm Optimization (PSO) into the neural network. Furthermore, Zhang et al. [49] proposed a novel ensemble method by taking AdaBoost, generic algorithm, and probabilistic SVM into consideration to acquire better prediction performance. Moreover, several hybrid based approaches have been proposed for stock trend prediction, which have achieved excellent performance. For example, Marković et al. [50] proposed a new hybrid method that integrates the analytic hierarchy process and weighted kernel least squares SVM. Lei [51] developed a hybrid method by combining the rough set and wavelet neural network. Many researchers have found that fusion model based on different techniques plays a vital role in prediction. Based on technical analysis and sentiment embedding, Picasso et al. [52] proposed a fusion model, which integrates Random Forest (RF), ANN, and SVM. Similar to [52], Parray et al. [53] applied several

methods based on machine learning including Logistic Regression (LR), SVM, and ANN to predict the trend of the stocks for the next day. Xu et al. [54] proposed a new fusion method by combining the k-mean clustering and ensemble method (i.e., SVM and RF). In order to reduce the influence of parameters, Dash et al. [55] proposed a new stock price prediction method named fine-tuned SVR, which combines the grid search technique and SVR. The grid search technique is used to select the best kernel function and tune the optimized parameters through training and validation datasets.

Apart from the aforementioned approaches based on machine learning, a variety of Deep Learning (DL) techniques have emerged in recent years for financial time series forecasting researches. In the next section, we will introduce some stock price prediction methods based on DL techniques.

2.2. Deep Learning Methods. DL technique based on ANNs is a branch of ML, which can extract high-level abstraction feature for data representation. DL methods can achieve excellent performance compared to conventional ML methods; thus, they have been widely applied in many application fields such as image processing and computer vision. Recently, DL methods have been proposed for analyzing the financial time series data. Specifically, combining the stock technical indicators and the historical price of stock data, Nelson et al. [20] applied LSTM model to predict the movement of the stock market's price. In order to forecast the short-term stock market trend, Liang et al. [56] constructed a new prediction model by combining Restricted Boltzmann Machine (RBM) and several classifiers. Kim and Won [57] combined LSTM with Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model to identify the stock price volatility. Moews et al. [58] designed a forecasting method based on Deep Feed-forward Neural Network (DFNN) and exponential smoothing. In order to reduce the training complexity and improve the prediction accuracy, Li et al. [59] constructed a forecasting model by integrating Feature Selection (FS) and LSTM method. For selecting and focusing on key information of stock data, Zhao et al. [60] introduced AM into RNN and proposed three prediction frameworks named AT-RRR, AT-LSTM, and AT-GRU, respectively.

CNN utilizes the local perception and weighted sharing to greatly reduce the number of parameters. Therefore, many methods based on CNN are recently proposed to predict the stock trend and achieve good performance. For example, Sezer and Ozbayoglu [61] first transformed stock technical indicators into 2D images and then designed a novel method based on CNN for the stock price prediction task. Wen et al. [62] first exploited sequences reconstruction method to reduce noise of the financial time series data. And then, they employed CNN model to extract spatial structure from the time series data for stock prediction. Different from [62], Barra et al. [63] first utilized Gramian angular field technique to obtain 2D images from the time series data. Then, an ensemble learning framework of CNNs was exploited to

forecast the trend of US market. Long et al. [64] defined three matrices to describe the trading behavior pattern named as the transaction number matrix, buying volume matrix, and selling volume matrix. Next, they exploited CNN to extract deep features. To capture the characteristics of time series, Hao and Gao [65] proposed a novel method by extracting multiscale CNN features of price series. Different from the existing methods, Lu et al. [66] proposed a new network model to predict stock price by combining CNN and LSTM.

3. Backgrounds

In this section, some basic knowledge of the proposed method will be reviewed briefly.

3.1. Long Short-Term Memory Network. RNN model is widely used to analyze and predict time series data [15]. However, RNN model often suffers from the vanishing gradient problem in the training processing. Therefore, it is difficult to remember the previous information, namely, long dependence problem [16, 67]. To deal with this issue, Greff et al. [67] have proposed a LSTM network model, which has the function of memory in a longer time span. The model utilizes gate control mechanism to adjust information flow and systematically determines the amount of incoming information to be retained in each time step. Figure 1 shows the structure of the basis LSTM unit composed of a storage unit and three control gates (named as input gate, output gate, and forgetting gate). x_t and h_t correspond to the input and hidden states of time t , respectively. f_t , i_t , and o_t are forgetting gate, input gate, and output gate, respectively. \tilde{C}_t is the candidate information for the input to be stored, and the amount of storage is then controlled by the input gate. The calculation processes of each gate, input candidate, cell state, and hidden state are shown in the following formulas:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\ \tilde{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \\ C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \\ h_t &= o_t \cdot \tanh(C_t), \end{aligned} \quad (1)$$

where W_f , W_i , W_o , and W_c represent the weight matrix of forgetting gate, input gate, output gate, and update state, respectively. b_f , b_i , b_o , and b_c represent the bias vectors of forgetting gate, input gate, output gate, and update state, respectively. x^t represents the time series data of current time interval t , and h_{t-1} is the output of memory unit in the previous time interval $t - 1$.

3.2. Attention Mechanism. Squeeze-and-Excitation (SENet) is an efficient attention mechanism, which can improve the representation ability of the network by modeling the dependency of each channel and can adjust the features

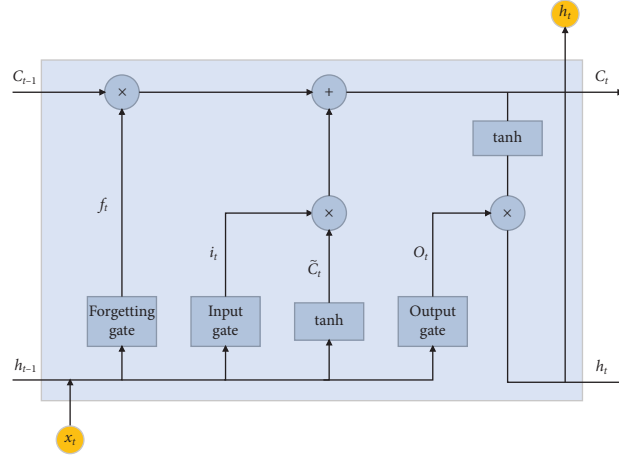


FIGURE 1: The structure diagram of LSTM unit.

channel by channel, so that the network can select more useful features [68]. The basic structure of SE block is shown in Figure 2. The first step is the squeeze operation, which takes the global spatial features of each channel as the representation of the channel to get the global feature description. The second step is the extraction operation, which learns the dependence degree of each channel and adjusts it, and the obtained feature graph is the output of SE block.

4. The Proposed Network Model Structure

In order to extract features efficiently and improve prediction accuracy, we combine CNN, BiLSTM network, and a lightweight ECA attention module into a unified framework and propose a new time series prediction network model named as CNN-BiLSTM-ECA. Our proposed model can automatically learn and extract local features and long memory features in the time series by making full use of data information to minimize the complexity of the model. The network model structure is shown in Figure 3.

In this model, firstly, the CNN model is utilized to extract deep feature vectors from the input origin time series data. Then, the BiLSTM model is employed to learn the temporal features from the new time series data constructed by the deep feature vectors. Moreover, the attention mechanism named as ECA is further introduced to extract more important features. Finally, the Dense model consisting of several fully connected layers is employed to perform the prediction task.

4.1. Convolutional Neural Network. In this paper, Convolution Neural Network (CNN) is utilized to extract data features efficiently. Compared with the traditional neural network structure, CNN is a local connection between neurons, which reduces the number of the parameters between the connection layers. In other words, it contains a part of the connections between $n-1$ layer and n layer in CNN. Figure 4 shows the difference between full connection and local connection, where Figure 4(a) is a schematic

diagram of a full connection and Figure 4(b) is a local connection diagram.

4.2. Bidirectional LSTM. In order to build a more accurate prediction model, the bidirectional LSTM (BiLSTM) network is employed, which acts as forward and backward LSTM networks for each training sequence. The two LSTM networks are connected to the same output layer to provide complete context information to each sequence point. Figure 5 shows the structure of BiLSTM.

4.3. Attention Module ECA. Channel Attention (CA) mechanism has great potential in improving the performance of Deep Convolutional Neural Networks (DCNNs). However, most of the existing methods are committed to designing more complex attention modules to achieve better performance, which inevitably increases the complexity and computational burden of the model. In order to avoid overfitting of the model and reduce the computation, a lightweight and low complexity module called as Effective Channel Attention (ECA) [32] is introduced. ECA can not only generate the weights for each channel, but also learn the correlation among the different channels. For the time series data, the larger weights will be assigned for the key features and smaller weights for the irrelevant features. Therefore, ECA focuses on the useful information, which improves the sensitivity of the network to the main features. Figure 6 shows the structure of ECA.

As shown in Figure 6, ECA is first to carry out channel Global Average Pooling (GAP). Then, ECA employs each channel and its k adjacent channels to capture the local cross channel interactions. ECA generates channel weights by performing fast 1D convolution as follows:

$$\omega = \sigma(\text{CID}_k(y)), \quad (2)$$

where CID is denoted as 1D convolution and k is the kernel size of 1D convolution.

In order to avoid manually adjusting k , ECA uses channel dimension adaptively mapping way to

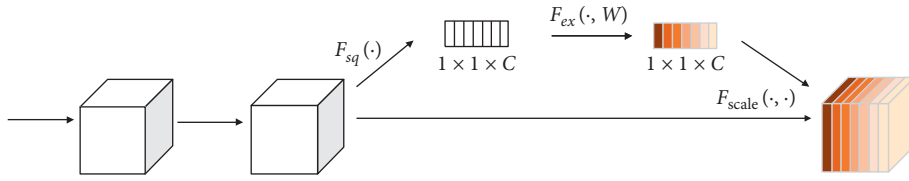


FIGURE 2: SE block structure diagram.

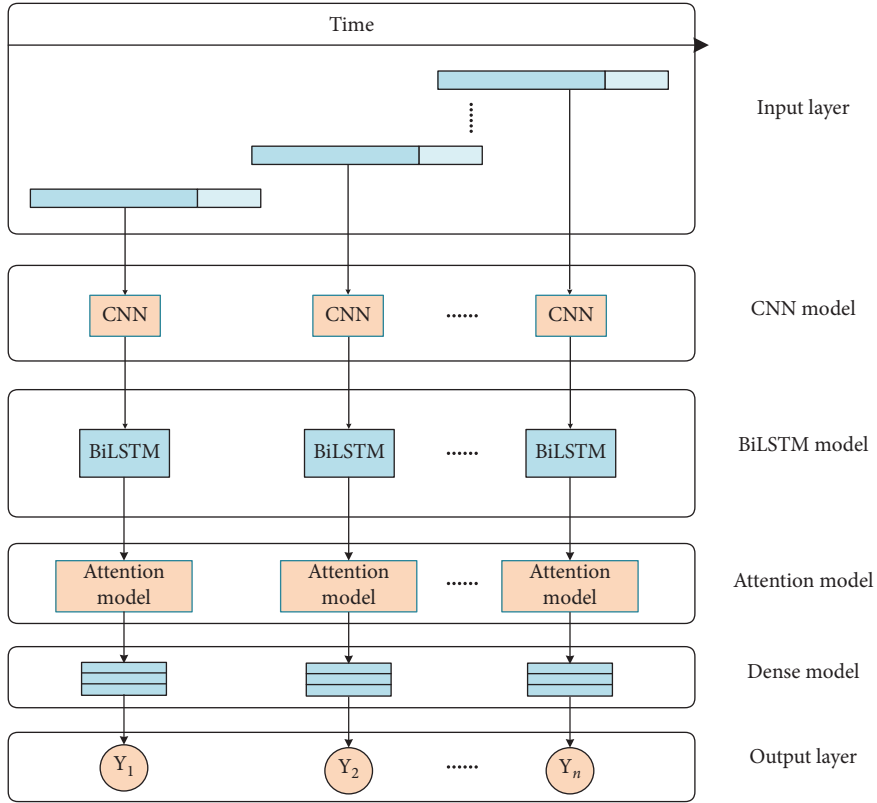


FIGURE 3: Network model structure diagram.

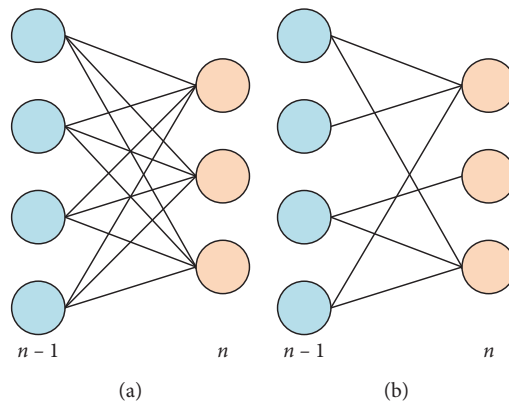


FIGURE 4: The diagram of full connection and local connection.

determine the value of k . Since the kernel size k of 1D convolution is directly proportional to the channel dimension C , the corresponding relationship is defined as follows:

$$C = \phi(k) = 2^{(\gamma * k - b)}. \tag{3}$$

Therefore, given the channel dimension C , the kernel size k can be adaptively determined by

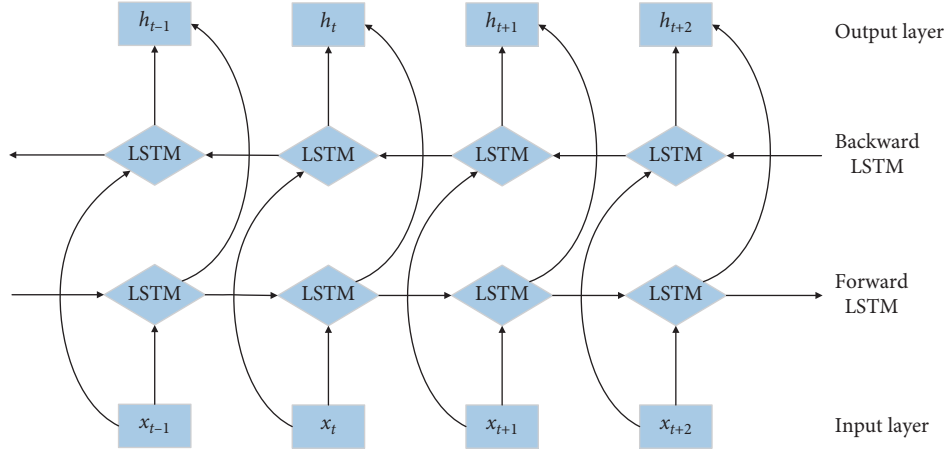


FIGURE 5: BiLSTM structure diagram.

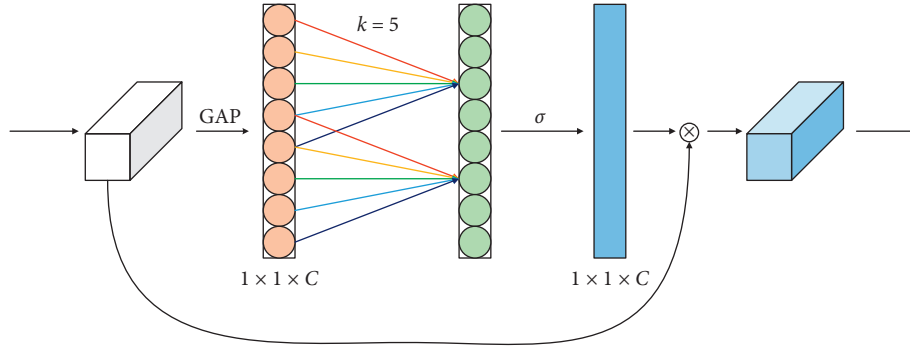


FIGURE 6: The structure diagram of ECA.

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma_{\text{odd}}} \right\rfloor, \quad (4)$$

where $\lfloor \cdot \rfloor_{\text{odd}}$ is the nearest odd number. The parameters of γ and b are set to 2 and 1, respectively, in this paper. Obviously, the high-dimensional channel has a longer range of interaction through nonlinear mapping, while the low dimensional channel has a shorter range of interaction.

4.4. Dense Model. The purpose of Dense model is first to exploit full connection layer to extract the correlation among these features changed by nonlinear mapping and then maps them to the output space. In the proposed network structure model, three layers of full connection layer are added to solve the nonlinear problem well, which can achieve accurate prediction. Figure 7 shows the structure of Dense model.

5. Experiments

5.1. Experimental Process. The experiment consists of six parts: data collecting, data preprocessing, model training, model saving, model testing, and prediction results. The

schematic diagram of the experiment process is shown in Figure 8.

5.2. Data Description and Preprocessing. The experimental data are collected from NetEase Finance (data collected from the website <http://quotes.money.163.com/0000001.html#1b01>), which includes Shanghai Composite Index (SSE Index for short, stock symbol: 000001), China Unicom (stock symbol: 600050), and CSI 300 (stock symbol: 399300). SSE stock data was collected from December 20, 1990, to November 23, 2020 (7304 groups); China Unicom stock data was collected from October 9, 2002, to March 17, 2021 (4340 groups); CSI 300 stock data was collected from January 7, 2002, to March 17, 2021 (4567 groups). Each stock data includes the closing price, the highest price, the lowest price, the opening price, previous day's closing price, change, ups and downs, and other time series data. Table 1 is an overview of the three stock's data, and Tables 2–4, respectively, show partial data of the three stocks.

From these tables, the original dataset obtained from the network cannot be directly used for model training and testing, so the data preprocessing is required. In view of the missing or disordered attribute values in the original dataset, we first use interpolation and sorting by date and other operations to complete time series data in our experiment. Then, data normalization is employed to deal with the

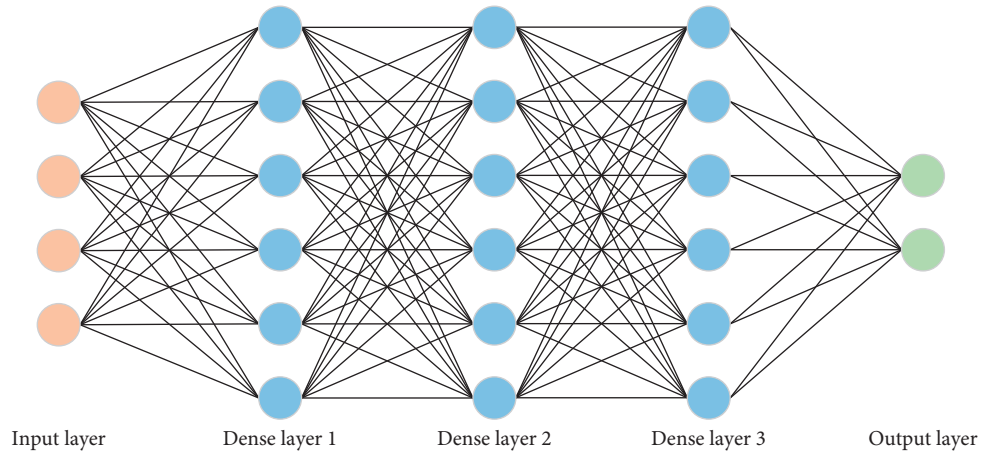


FIGURE 7: The structure diagram of Dense model.

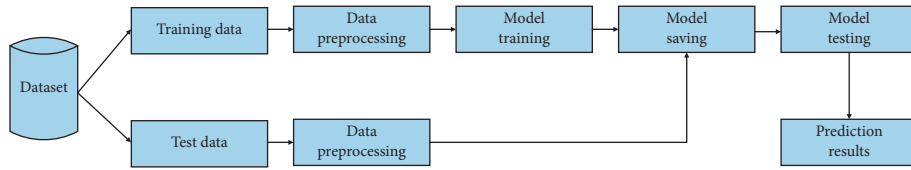


FIGURE 8: The flow chart of the experiment.

TABLE 1: The detailed information of three stocks' data.

Stock code	Stock name	Start and end time	Total amount of data/group
000001	SSE	1990.12.20–2020.11.23	7304
600050	China Unicom	2002.10.09–2021.03.17	4340
399300	CSI 300	2002.01.07–2021.03.17	4657

TABLE 2: Some data come from SSE stock.

Data	Closing price	Highest price	Lowest price	Opening price	Previous day's closing price	Change	Ups and downs
1990/12/20	104.39	104.39	99.98	104.3	99.98	4.41	4.4109
1990/12/21	109.13	109.13	103.73	109.07	104.39	4.74	4.5407
1990/12/24	114.55	114.55	109.13	113.57	109.13	5.42	4.9666
1990/12/25	120.25	120.25	114.55	120.09	114.55	5.7	4.976
1990/12/26	125.27	125.27	120.25	125.27	120.25	5.02	4.1746
1990/12/27	125.28	125.28	125.27	125.27	125.27	0.01	0.008
1990/12/28	126.45	126.45	125.28	126.39	125.28	1.17	0.9339
1990/12/31	127.61	127.61	126.48	126.56	126.45	1.16	0.9174

TABLE 3: Some data come from China Unicom stock.

Data	Closing price	Highest price	Lowest price	Opening price	Previous day's closing price	Change	Ups and downs
2002/10/09	2.87	3.15	2.86	3.05	2.3	0.57	24.7826
2002/10/10	2.83	2.89	2.79	2.83	2.87	-0.04	-1.3937
2002/10/11	2.94	2.96	2.83	2.85	2.83	0.11	3.8869
2002/10/14	3.01	3.02	2.94	2.96	2.94	0.07	2.381
2002/10/15	2.98	3.01	2.96	3.01	3.01	-0.03	-0.9967
2002/10/16	2.99	3.05	2.98	2.99	2.98	0.01	0.3356
2002/10/17	2.98	3.01	2.97	2.99	2.99	-0.01	-0.3344
2002/10/18	3.01	3.03	2.97	2.98	2.98	0.03	1.0067

TABLE 4: Some data come from CSI 300 stock.

Data	Closing price	Highest price	Lowest price	Opening price	Previous day's closing price	Change	Ups and downs
2002/1/7	1302.08	1302.08	1302.08	1302.08	1316.46	-14.38	-1.0923
2002/1/8	1292.71	1292.71	1292.71	1292.71	1302.08	-9.37	-0.7196
2002/1/9	1272.65	1272.65	1272.65	1272.65	1292.71	-20.06	-1.5518
2002/1/10	1281.26	1281.26	1281.26	1281.26	1272.65	8.61	0.6765
2002/1/11	1249.81	1249.81	1249.81	1249.81	1281.26	-31.45	-2.4546
2002/1/14	1205.15	1205.15	1205.15	1205.15	1249.81	-44.66	-3.5733
2002/1/15	1186.43	1186.43	1186.43	1186.43	1205.15	-18.72	-1.5533
2002/1/16	1201.88	1201.88	1201.88	1201.88	1186.43	15.45	1.3022

problem of inconsistency in the magnitude of the data. Finally, an effective dataset can be constructed for the experiment. Data normalization is to scale the data according to a certain scale and transform it into a specific interval. In this experiment, the value of data is converted to $[0, 1]$, which improves the convergence speed and accuracy of the model. The transformation function is defined as follows:

$$x^* = \frac{x - \min}{\max - \min}, \quad (5)$$

where max and min are denoted as the maximum value and minimum value of the sample data, respectively.

5.3. Experiment Setting and Implementation. Many hyper-parameters are required to be predefined, e.g., the neurons numbers of different layers, the learning rate, and the number of iterations. In our experiment, the neurons numbers of three fully connected layers in Dense model are set as 128, 32, and 1, respectively. The hidden neuron number of each layer in BiLSTM is all set to 64. The learning rate is 0.001, and the number of iterations is 200. Relu function is used as the activation function. Description of the parameter setting of the proposed method is shown in Table 5.

In this experiment, Keras is used as the framework of neural network and Python programming language is employed to implement network structure. According to the parameter setting of the proposed network, the specific CNN-BiLSTM-ECA model structure is shown in Figure 9.

5.4. Evaluation Criterion. In order to analyze the performance of the proposed prediction model intuitively and quantitatively, three evaluation criteria including Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are utilized to evaluate the prediction results comprehensively.

The Mean Square Error (MSE) represents the mean of the sum squares of the errors between the predicted data and the original data, which is defined as follows:

TABLE 5: Parameters setting of the proposed method.

	Description of the parameters	Value
CNN	The filter number of convolution layer in CNN	1
	The kernel_size of the convolution layer in CNN	1
	The activation function in CNN	Relu
	The way of convolution layer padding in CNN	Same
BiLSTM	The number of hidden units of BiLSTM	64
	The activation function of BiLSTM	Relu
Dense	The layers number of the Dense model	3
	The hidden neurons number of the first Dense layer	128
	The hidden neurons number of the second Dense layer	32
	The hidden neurons number of the third Dense layer	1
	The activation function of Dense model	Relu
Other	Time step	10
	Batch size	256
	Epochs	200
	Optimizer	Adam
	Loss function	MSE
	Learning rate	0.001

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n (X_t - X'_t)^2, \quad (6)$$

where X_t and X'_t represent the predicted value and the true value, respectively.

The Root Mean Square Error (RMSE) defines the expected value of the square of the error caused by the predicted value and the true value, and its range is $[0, +\infty)$. When the predicted value is completely consistent with the true value, the value of RMSE is equal to 0, which is a perfect model. On the contrary, when the prediction error is larger, the value of the RMSE will be larger. The calculation formula is as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (X_t - X'_t)^2}. \quad (7)$$

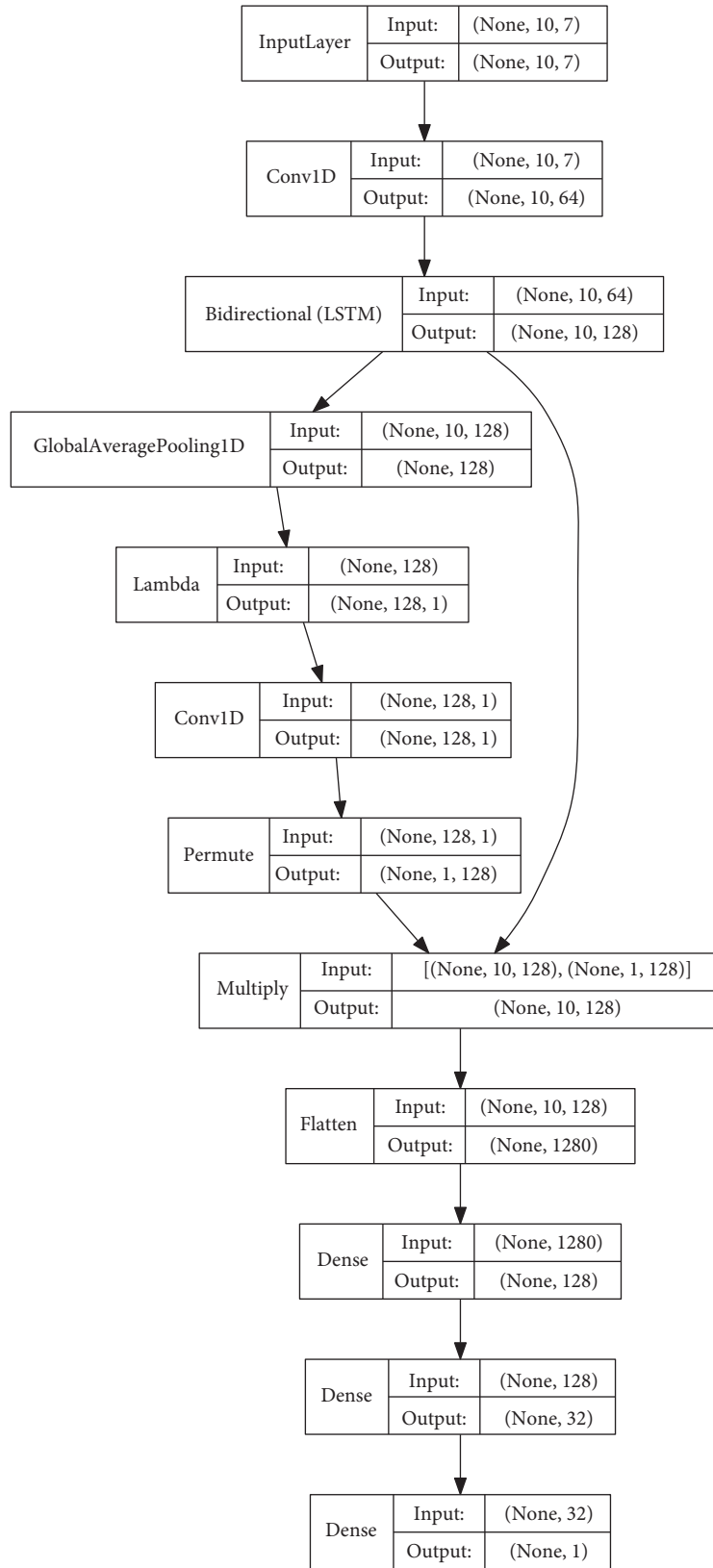


FIGURE 9: The model structure of the proposed CNN-BiLSTM-ECA using Keras to implement.

The mean absolute error (MAE) represents the average of the absolute value of the deviation of all single observations, which can avoid the problem of mutual cancellation of errors and accurately reflect the size of the actual forecast error as below:

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |X_t - X'_t|. \quad (8)$$

5.5. Experiment Results and Analysis

5.5.1. Parameter Analysis. Time step is a key parameter of the time series neural network, which affects the prediction performance of the network. Therefore, we analyze the time step parameters on the three datasets including the Shanghai Composite Index, China Mobile, and CSI 300. Firstly, the top 85% of the dataset is used for the training set, and the remaining 15% is used for the test set. Secondly, we conduct comparative experiments under the conditions that when the time steps are set as 5, 10, 15, and 20. The experimental results under different parameters are shown in Figures 10–12. The prediction error values of each time step are shown in Table 6.

From these figures, when the time step is set to 5, the prediction results have a large deviation. It means that the small-time step setting will ignore the influence of global factors. When the time step is set to 20, the time setting span is too large. Therefore, the data have certain fluctuations resulting in inaccurate prediction results. When the time step is set to 10, the error of the model is the smallest and the accuracy is the highest. Therefore, the time step is set to 10; in other words, the stock data of the previous 10 days is used as the input layer of the neural unit, and the closing price of the 11th day is used as the label to train the model.

Table 6 visually analyzes and describes the forecast error value of each time step. It is found that the error trend between each time step is consistent with the error trend between the predicted price curve and the actual price curve in Figures 10–12. Therefore, according to these results, when the time step is set to 10, the prediction model can achieve the best performance.

5.5.2. Experimental Results and Analysis. In order to verify the effectiveness of the proposed model, different models including CNN [10], LSTM [15], BiLSTM [25, 26],

CNN-LSTM [20], BiLSTM-ECA, CNN-LSTM-ECA, and CNN-BiLSTM [27] are compared on the three stock datasets collected from the Shanghai Composite Index, China Unicom, and CSI 300. The prediction results are shown in Figures 13–15. As shown in these figures, the blue curve is the predicted value of the closing price of the stock, and the red curve is the true value of the closing price of the stock. Meanwhile, the x -axis is the time, and the y -axis is the normalization value of the stock price.

As shown in Figures 13–15, firstly, we can find that the prediction results of BiLSTM and LSTM methods are better than CNN model. It indicates that LSTM and BiLSTM take the trend of the stock price as time series in consideration, which is useful to improve the accuracy of the forecast. Moreover, the performance of BiLSTM is better than LSTM due to the fact that BiLSTM model can exploit the subsequent information of price stock time series. Secondly, introducing the CNN model to reduce the noise and capture nonlinear structure of the stock data well, CNN-LSTM and CNN-BiLSTM methods can outperform CNN, LSTM, and BiLSTM. Finally, by integrating the ECA model into LSTM and BiLSTM to select the important features and key information, BiLSTM-ECA, CNN-LSTM-ECA, and CNN-BiLSTM-ECA methods can achieve better performance than other methods. To sum up, the proposed CNN-BiLSTM-ECA model obtains the best prediction results.

The prediction errors of different methods on the three stock datasets are listed in Table 7. As shown in this table, firstly, the CNN model has the low prediction performance among all the compared approaches. Since BiLSTM uses bidirectional information, BiLSTM model can further improve more prediction ability than LSTM. Then, owing to the deep feature extraction of CNN, CNN-LSTM and CNN-BiLSTM model have higher predictive power. At last, among all the compared methods, the CNN-BiLSTM-ECA model can achieve the best performance, which verifies the effectiveness of the CNN-BiLSTM-ECA network model.

In order to show the model performance more clearly, some prediction results of the LSTM model and the CNN-BiLSTM-ECA model on the part days of the three datasets are selected, as shown in Figures 16–18. From these figures, we can learn that the prediction results of the CNN-BiLSTM-ECA model are closer to the true value of the stock price, and the error is smaller. In summary, the experiment results verify the feasibility and effectiveness of the proposed network model.

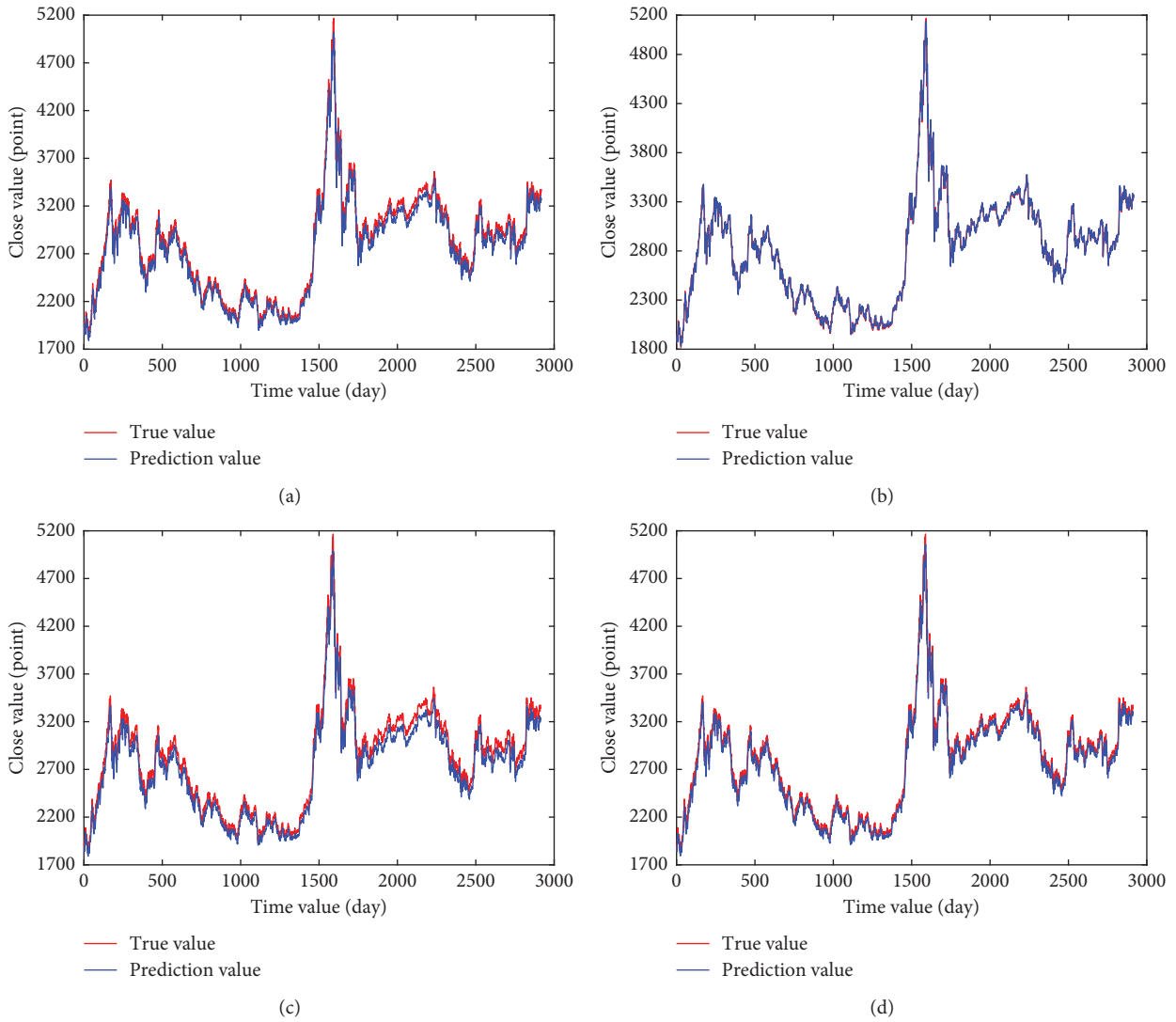


FIGURE 10: Comparison of the experimental results of different time steps of the Shanghai Composite Index. (a) Time step = 5. (b) Time step = 10. (c) Time step = 15. (d) Time step = 20.

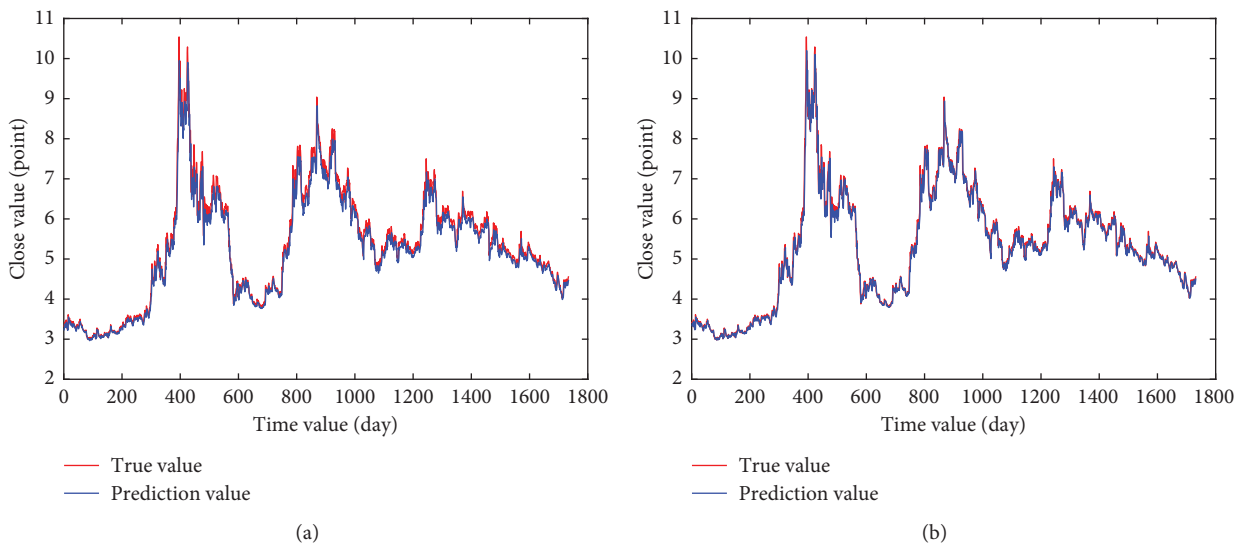


FIGURE 11: Continued.

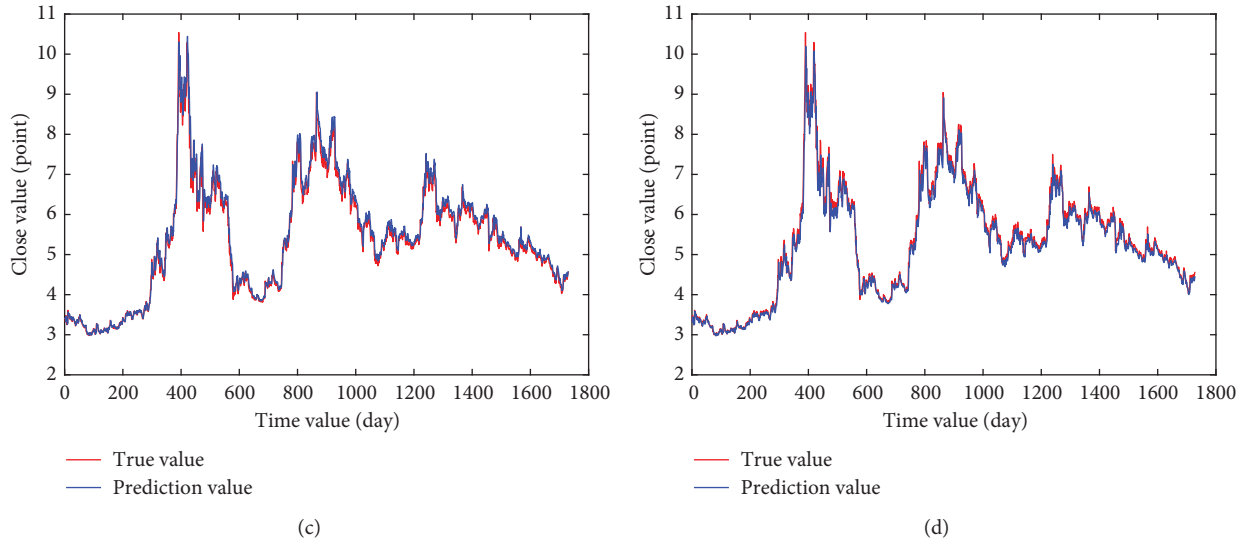


FIGURE 11: Comparison experiment results of different time steps of China Unicom. (a) Time step = 5. (b) Time step = 10. (c) Time step = 15. (d) Time step = 20.

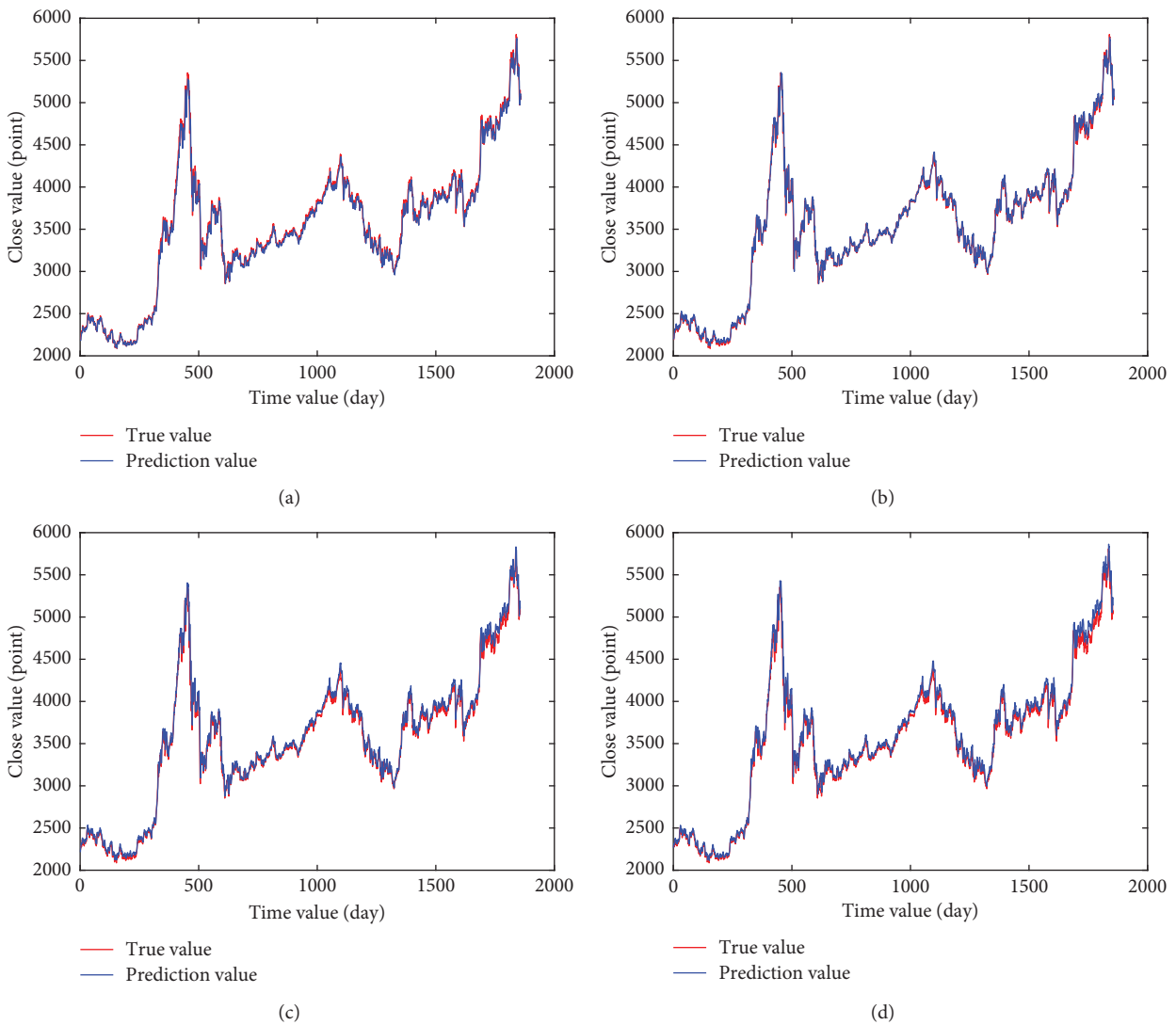


FIGURE 12: Comparison of experimental results of different time steps of CSI 300. (a) Time step = 5. (b) Time step = 10. (c) Time step = 15. (d) Time step = 20.

TABLE 6: Forecast errors at different time steps.

Time step	Shanghai Composite Index			China Unicom			CSI 300		
	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
5	6431.222	80.195	70.004	0.044	0.209	0.144	3949.252	62.843	43.987
10	1956.036	44.227	28.349	0.028	0.167	0.103	3434.408	58.604	39.111
15	10162.775	100.811	91.489	0.038	0.196	0.177	5013.827	70.808	51.035
20	4711.037	68.637	58.377	0.035	0.186	0.123	6409.845	80.062	57.951

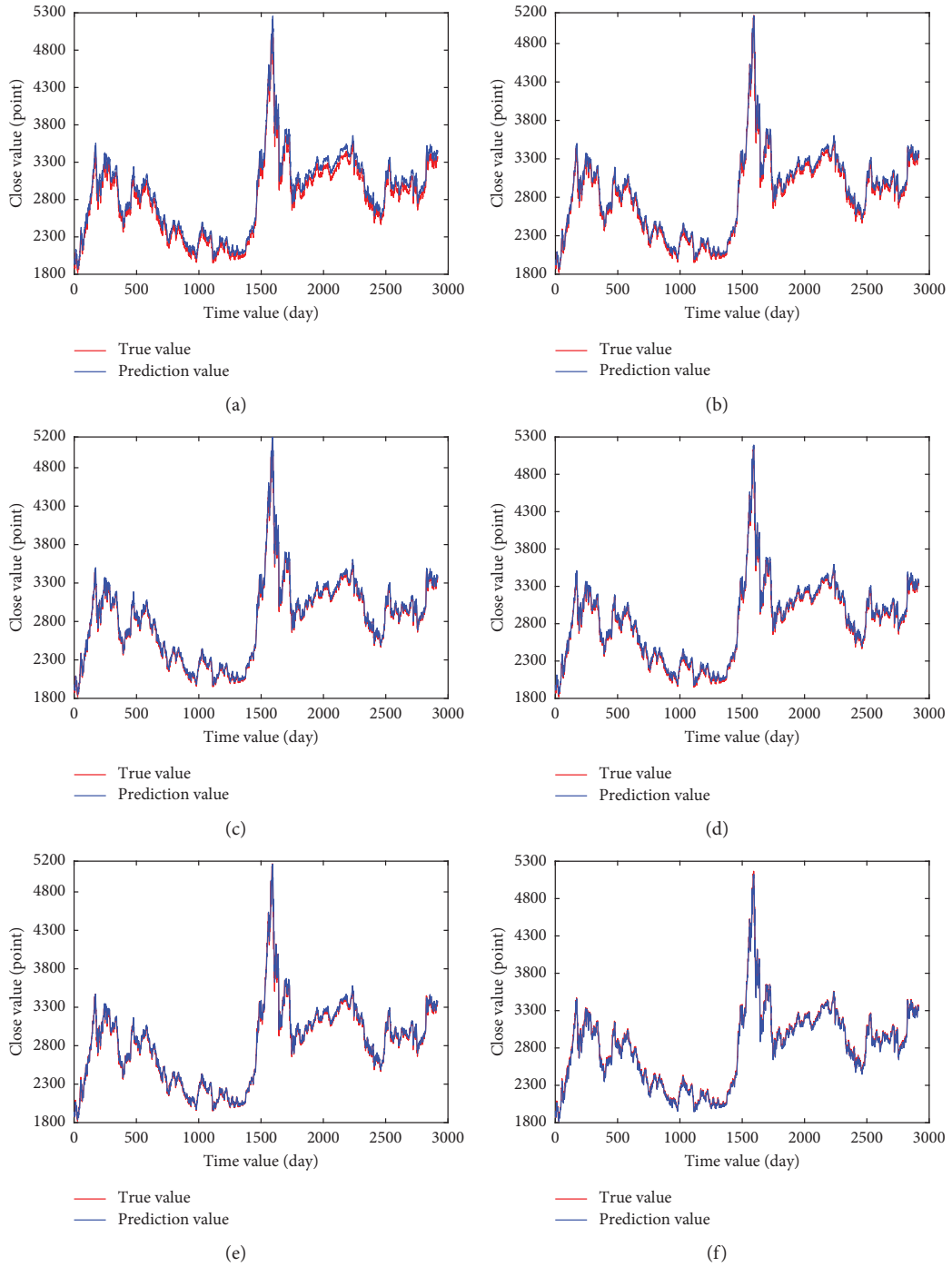


FIGURE 13: Continued.

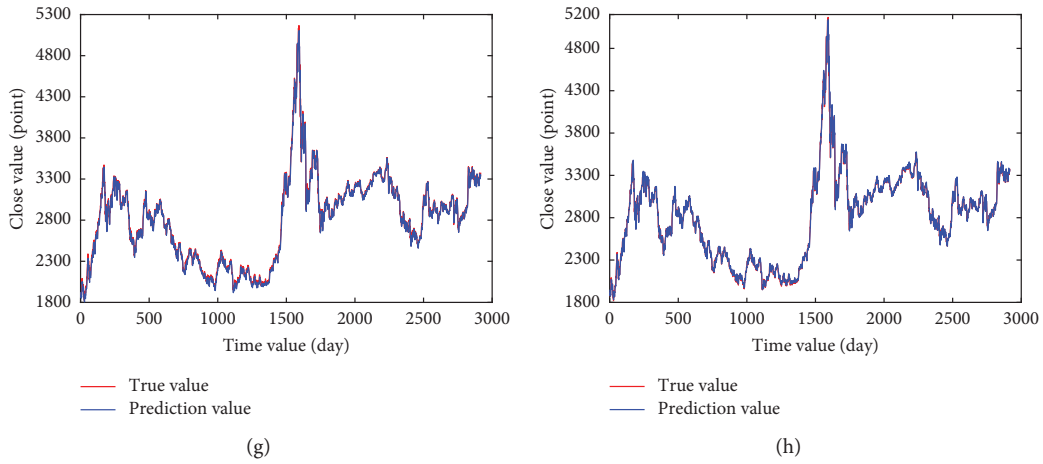


FIGURE 13: The prediction results of different models on the Shanghai Composite Index. (a) CNN. (b) LSTM. (c) BiLSTM. (d) CNN-LSTM. (e) CNN-BiLSTM. (f) BiLSTM-ECA. (g) CNN-LSTM-ECA. (h) CNN-BiLSTM-ECA.

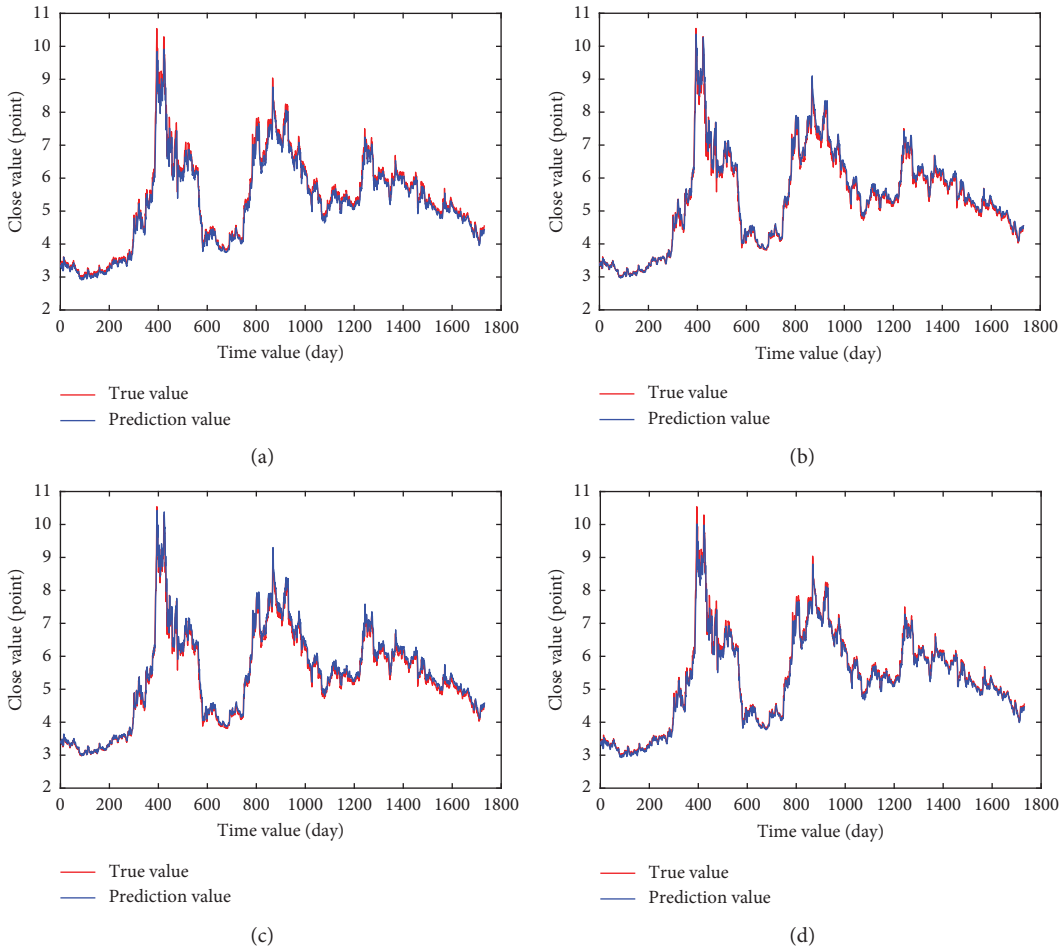


FIGURE 14: Continued.

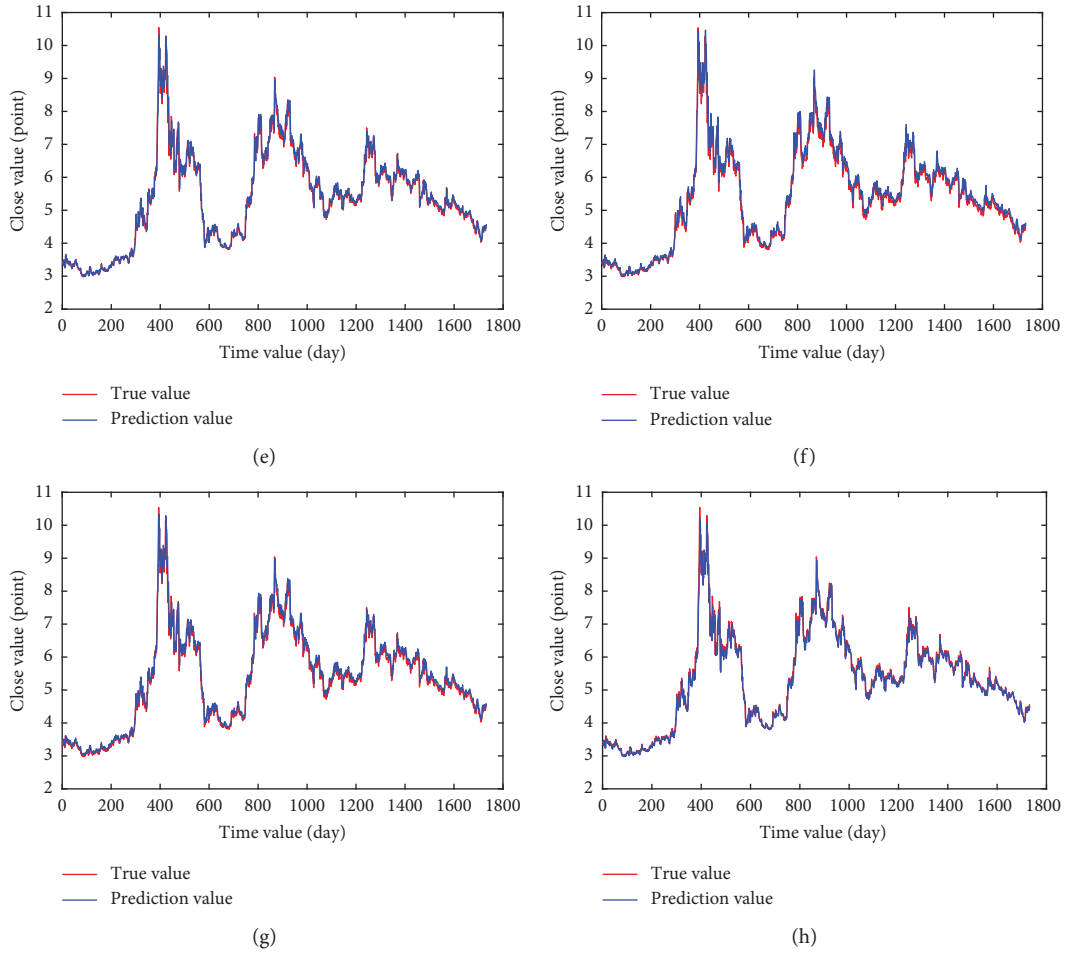


FIGURE 14: The prediction results of different models for China Unicom. (a) CNN. (b) LSTM. (c) BiLSTM. (d) CNN-LSTM. (e) CNN-BiLSTM. (f) BiLSTM-ECA. (g) CNN-LSTM-ECA. (h) CNN-BiLSTM-ECA.

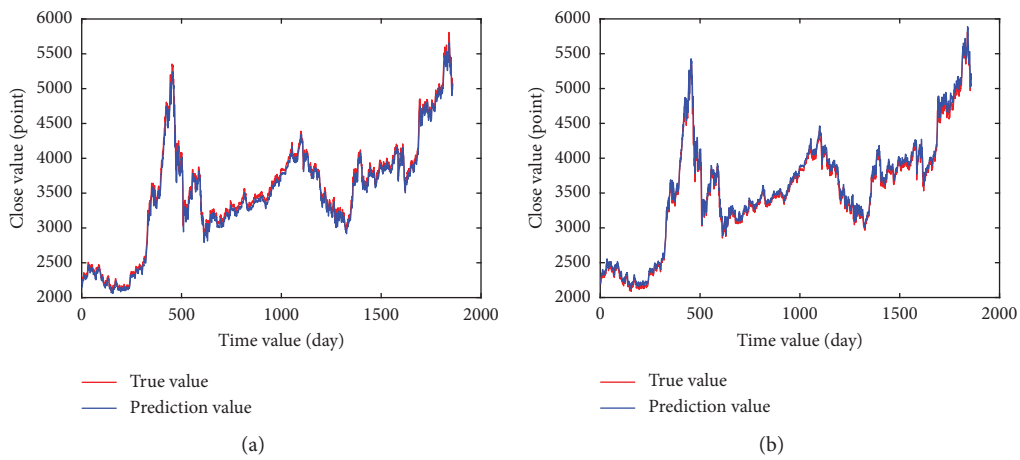


FIGURE 15: Continued.

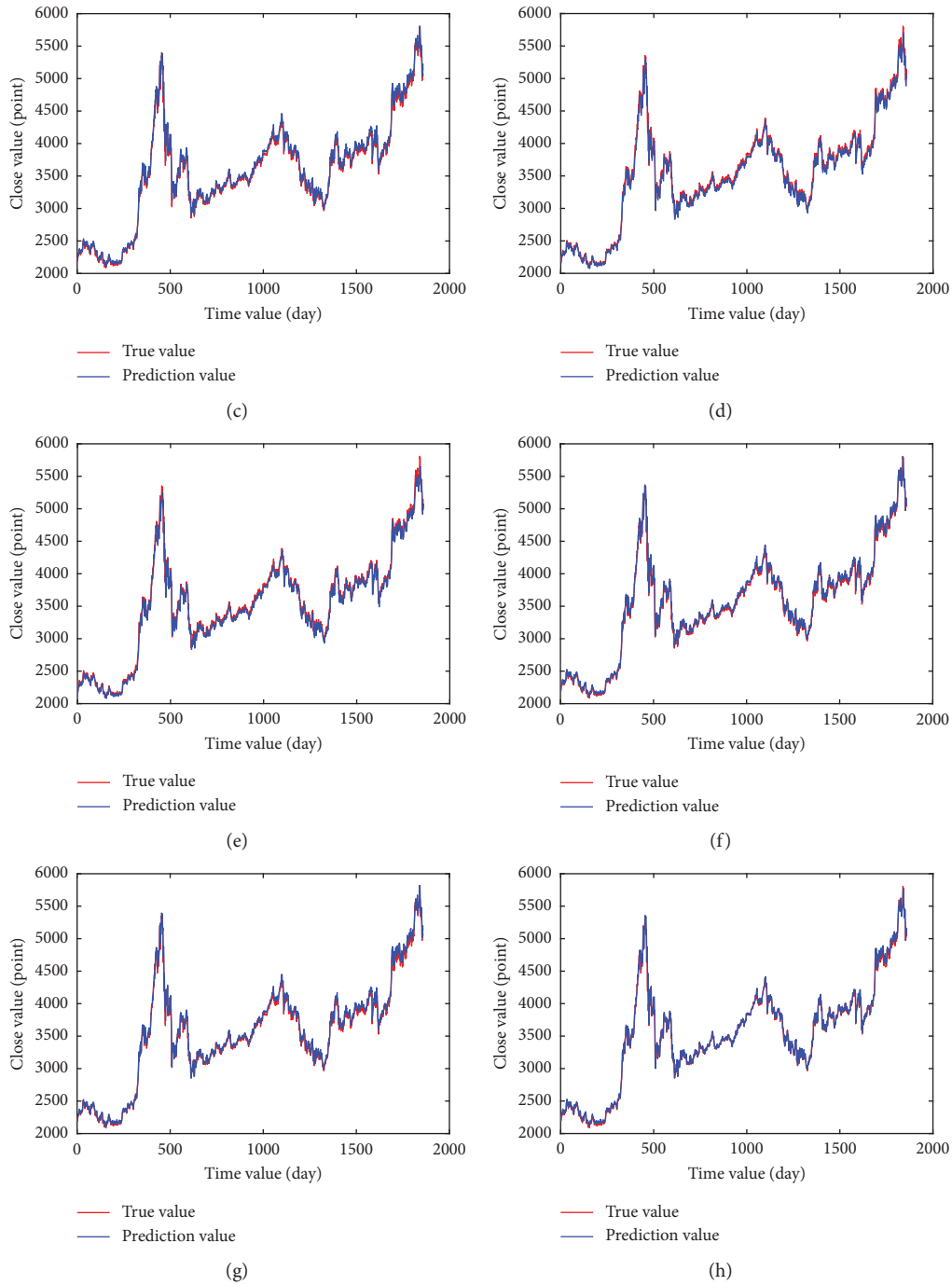


FIGURE 15: The prediction results of different models for CSI 300. (a) CNN. (b) LSTM. (c) BiLSTM. (d) CNN-LSTM. (e) CNN-BiLSTM. (f) BiLSTM-ECA. (g) CNN-LSTM-ECA. (h) CNN-BiLSTM-ECA.

TABLE 7: Forecast errors of different network models.

Model	Shanghai Composite Index			China Unicom			CSI 300		
	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
CNN	8447.149	91.908	79.914	0.037	0.193	0.134	6218.092	78.855	63.981
LSTM	4222.102	64.978	50.585	0.036	0.189	0.128	5809.153	76.218	58.679
BiLSTM	2603.726	51.027	35.419	0.035	0.187	0.132	5091.610	71.356	52.119
CNN-LSTM	2902.648	53.876	39.502	0.030	0.174	0.110	4905.472	70.039	52.457
CNN-BiLSTM	2321.235	48.179	32.289	0.029	0.170	0.110	4643.541	68.144	51.143
BiLSTM-ECA	2184.278	46.736	32.737	0.039	0.198	0.142	4161.203	64.507	46.453
CNN-LSTM-ECA	2200.705	46.911	32.353	0.032	0.180	0.127	4568.808	67.593	51.061
CNN-BiLSTM-ECA	1956.036	44.227	28.349	0.028	0.167	0.103	3434.408	58.604	39.111

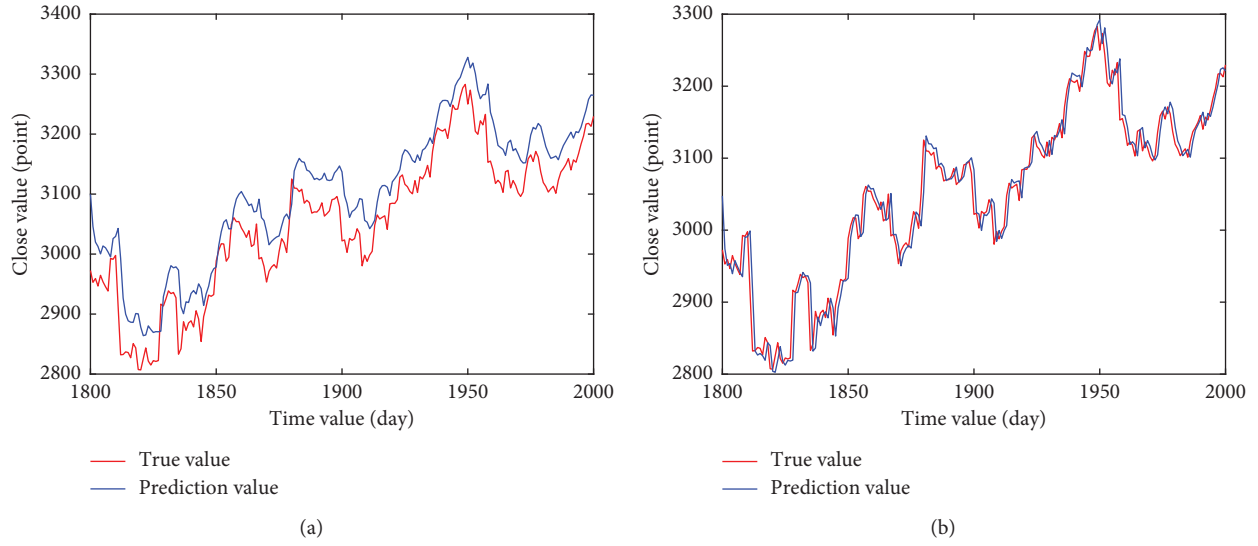


FIGURE 16: Local prediction results of LSTM and CNN-BiLSTM-ECA models on the Shanghai composite index. (a) LSTM. (b) CNN-BiLSTM-ECA.

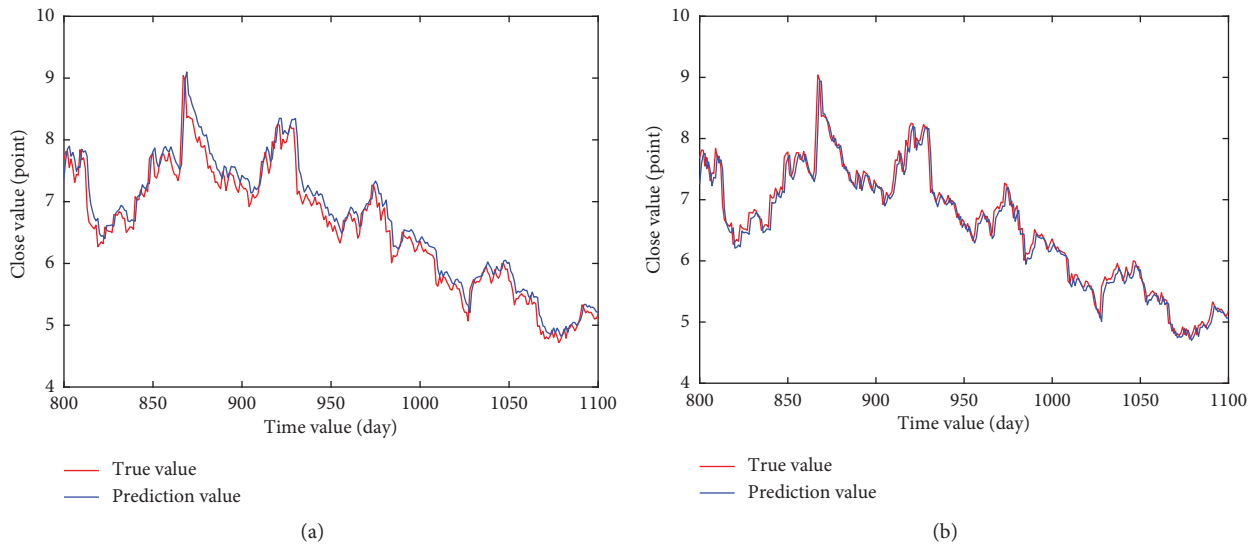


FIGURE 17: Local prediction results of LSTM and CNN-BiLSTM-ECA models on the China Unicom. (a) LSTM. (b) CNN-BiLSTM-ECA.

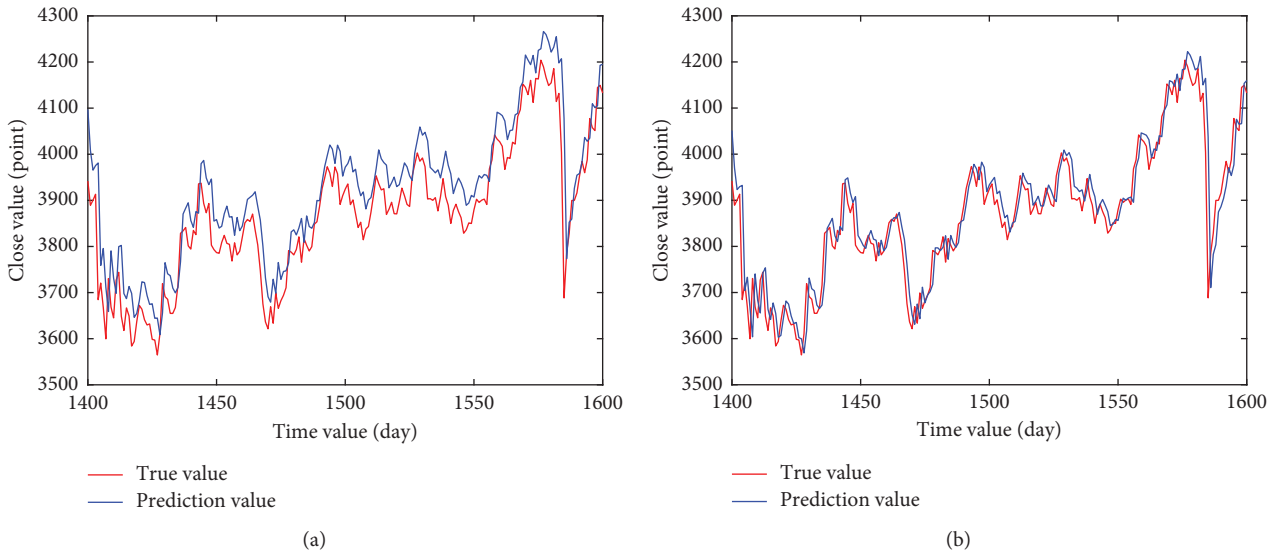


FIGURE 18: Local prediction results of LSTM and CNN-BiLSTM-ECA models on the CSI 300. (a) LSTM. (b) CNN-BiLSTM-ECA.

6. Conclusions

For the stock market, learning the future price is very important for making investment decisions. This paper first proposes a new stock price time series forecasting network model (CNN-BiLSTM-ECA), which takes the stock closing price, the highest price, the lowest price, the opening price, the previous day's closing price, the change, the rise and fall, and other time series data as the input to predict the next day's stock closing price. The proposed network model combines CNN and BiLSTM network models. Firstly, CNN is used to extract the deep features of the input data effectively. Secondly, feature vectors are constructed in time series as inputs to the BiLSTM network for learning and prediction. At the same time, ECA attention model is introduced into the model to enhance the importance of learning features. Thus, the useless features are reduced and the accuracy of the model is further improved. The proposed model is compared with CNN, LSTM, BiLSTM, CNN-LSTM, CNN-BiLSTM, BiLSTM-ECA, and CNN-LSTM-ECA network models on three datasets. The experimental results show that the proposed model has the highest prediction accuracy and the best performance. The MSE, RMSE, and MAE of CNN-BiLSTM-ECA are the smallest among all methods, with the values of 1956.036, 44.227, and 28.349, respectively, for the Shanghai Composite Index. Compared with single LSTM model, the reduction was 53.67%, 31.94%, and 43.96%, respectively. It can be seen that it is difficult to achieve high prediction accuracy by using only a single network, but the complexity of the network can improve its prediction accuracy. The CNN-BiLSTM-ECA model proposed in this paper can effectively predict stock prices and provide relevant references for investors to maximize investment returns.

The existing problem is that the input characteristic parameters of the model selection are relatively single. Therefore, it can be considered to increase the training features of the model, such as the emotional or sentiment features of news events or social media [69], so as to improve the prediction performance of the model from the perspective of feature selection. And then, we further apply our model on more application fields, such as gold price prediction, oil price prediction, foreign exchange price prediction, novelty detection [70], and optic disc detection [71]. In addition, a graph-based embedding technology will be introduced to solve the problem of time series prediction [72].

Data Availability

The network code and data are available from the corresponding author upon request.

Conflicts of Interest

All the authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62062040 and 61967010), Provincial Key Research and Development Program of Jiangxi (20192BAB217007), the China Postdoctoral Science Foundation (2019M661117), the Scientific Research Fund Project of Liaoning Provincial Department of Education (JYT19040 and JYT19053), the Natural Science Foundation of Liaoning Province Science and Technology Department (2019-ZD-0234), the Young Talent Cultivation Program of Jiangxi Normal University, Graduate Innovation Project of Jiangxi Provincial Department of Education (YC2020-S177), and the Graduate Innovation Foundation Project of Jiangxi Normal University (YJS2020045).

References

- [1] Y. Yi, J. Wang, W. Zhou, C. Zheng, J. Kong, and S. Qiao, "Non-negative matrix factorization with locality constrained adaptive graph," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 2, pp. 427–441, 2020.
- [2] Y. Yi, J. Wang, W. Zhou, Y. Fang, J. Kong, and Y. Lu, "Joint graph optimization and projection learning for dimensionality reduction," *Pattern Recognition*, vol. 92, pp. 258–273, 2019.
- [3] R. C. Cavalcante, R. C. Brasileiro, V. L. F. Souza, J. P. Nobrega, and A. L. I. Oliveira, "Computational intelligence and financial markets: a survey and future directions," *Expert Systems with Applications*, vol. 55, pp. 194–211, 2016.
- [4] H. Song, D. Peng, and X. Huang, "Incorporating research reports and market sentiment for stock excess return prediction: a case of mainland china," *Scientific Programming*, vol. 2020, Article ID 8894757, 7 pages, 2020.
- [5] A. Thakkar and K. Chaudhari, "Fusion in stock market prediction: a decade survey on the necessity, recent developments, and potential future directions," *Information Fusion*, vol. 65, pp. 95–107, 2021.
- [6] G. Kumar, S. Jain, and U. P. Singh, "Stock market forecasting using computational intelligence: a survey," *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1069–1101, 2021.
- [7] I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A systematic review of fundamental and technical analysis of stock market predictions," *Artificial Intelligence Review*, vol. 53, no. 4, pp. 3007–3057, 2020.
- [8] L. De Haan, C. Mercadier, and C. Zhou, "Adapting extreme value statistics to financial time series: dealing with bias and serial dependence," *Finance and Stochastics*, vol. 20, no. 2, pp. 321–354, 2016.
- [9] H. Tang, P. Dong, and Y. Shi, "A new approach of integrating piecewise linear representation and weighted support vector machine for forecasting stock turning points," *Applied Soft Computing*, vol. 78, pp. 685–696, 2019.
- [10] W. Chen, M. Jiang, W.-G. Zhang, and Z. Chen, "A novel graph convolutional feature based convolutional neural network for stock trend prediction," *Information Sciences*, vol. 556, pp. 67–94, 2021.
- [11] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: a systematic

- literature review: 2005–2019,” *Applied Soft Computing*, vol. 90, Article ID 106181, 2020.
- [12] P. Yu and X. Yan, “Stock price prediction based on deep neural networks,” *Neural Computing and Applications*, vol. 32, no. 6, pp. 1609–1628, 2020.
 - [13] E. Hajizadeh, A. Seifi, M. H. Fazel Zarandi, and I. B. Turksen, “A hybrid modeling approach for forecasting the volatility of S&P 500 index return,” *Expert Systems with Applications*, vol. 39, no. 1, pp. 431–436, 2012.
 - [14] A. M. Rather, A. Agarwal, and V. N. Sastry, “Recurrent neural network and a hybrid model for prediction of stock returns,” *Expert Systems with Applications*, vol. 42, no. 6, pp. 3234–3241, 2015.
 - [15] J. T. Connor, R. D. Martin, and L. E. Atlas, “Recurrent neural networks and robust time series prediction,” *IEEE transactions on Neural Networks*, vol. 5, no. 2, pp. 240–254, 1994.
 - [16] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, “Deep learning with long short-term memory for time series prediction,” *IEEE Communications Magazine*, vol. 57, no. 6, pp. 114–119, 2019.
 - [17] G. Ding and L. Qin, “Study on the prediction of stock price based on the associated network model of LSTM,” *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 6, pp. 1307–1317, 2020.
 - [18] X. Yan, W. Weihang, and M. Chang, “Research on financial assets transaction prediction model based on LSTM neural network,” *Neural Computing and Applications*, vol. 33, no. 1, pp. 257–270, 2021.
 - [19] Y. Baek and H. Y. Kim, “ModAugNet: a new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module,” *Expert Systems with Applications*, vol. 113, pp. 457–480, 2018.
 - [20] D. M. Q. Nelson, A. C. M. Pereira, and R. A. de Oliveira, “Stock market’s price movement prediction with LSTM neural networks,” in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 1419–1426, Anchorage, AK, USA, 2017.
 - [21] P.-Y. Hao, C.-F. Kung, C.-Y. Chang, and J.-B. Ou, “Predicting stock price trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane,” *Applied Soft Computing*, vol. 98, Article ID 106806, 2021.
 - [22] C.-H. Cheng, T.-L. Chen, and L.-Y. Wei, “A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting,” *Information Sciences*, vol. 180, no. 9, pp. 1610–1629, 2010.
 - [23] Y. Zhang and L. Wu, “Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network,” *Expert Systems with Applications*, vol. 36, no. 5, pp. 8849–8854, 2009.
 - [24] D. Zhang and S. Lou, “The application research of neural network and BP algorithm in stock price pattern classification and prediction,” *Future Generation Computer Systems*, vol. 115, pp. 872–879, 2021.
 - [25] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, 2005.
 - [26] Y. Xu, L. Chhim, B. Zheng, and Y. Nojima, “Stacked deep learning structure with bidirectional long-short term memory for stock market prediction,” in *Proceedings of the 2020 International Conference on Neural Computing for Advanced Applications*, pp. 447–460, Shenzhen, China, 2020.
 - [27] Q. Chen, W. Zhang, and Y. Lou, “Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long-short term memory neural network,” *IEEE Access*, vol. 8, pp. 117365–117376, 2020.
 - [28] W. Lu, J. Li, J. Wang, and L. Qin, “A CNN-BiLSTM-AM method for stock price prediction,” *Neural Computing and Applications*, vol. 33, pp. 4741–4753, 2020.
 - [29] W. Chen, S. Wang, X. Zhang et al., “EEG-based motion intention recognition via multi-task RNNs,” in *Proceedings of the 2018 SIAM International Conference on Data Mining*, San Diego, CA, USA, 2018.
 - [30] W. Chen, S. Wang, G. Long, L. Yao, Q. Z. Sheng, and X. Li, “Dynamic illness severity prediction via multi-task RNNs for intensive care unit,” in *Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM)*, IEEE, Singapore, 2018.
 - [31] S. Kastner and L. G. Ungerleider, “Mechanisms of visual attention in the human cortex,” *Annual Review of Neuroscience*, vol. 23, no. 1, pp. 315–341, 2000.
 - [32] Z. Yang, X. He, J. Gao, L. Deng, and A. Smola, “Stacked attention networks for image question answering,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 2016.
 - [33] Z. Yu, J. Yu, Y. Cui, D. Tao, and Q. Tian, “Deep modular co-attention networks for visual question answering,” in *Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019.
 - [34] H. Choi, K. Cho, and Y. Bengio, “Fine-grained attention mechanism for neural machine translation,” *Neurocomputing*, vol. 284, pp. 171–176, 2018.
 - [35] A. Galassi, M. Lippi, and P. Torrioni, “Attention in natural language processing,” *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–18, 2020.
 - [36] T. Afouras, J. S. Chung, A. Senior, O. Vinyals, and A. Zisserman, “Deep audio-visual speech recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–11, 2018.
 - [37] C. Li and Y. Qian, “Deep audio-visual speech separation with attention mechanism,” in *Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7314–7318, Barcelona, Spain, 2020.
 - [38] D. T. Tran, A. Iosifidis, J. Kannianen, and M. Gabbouj, “Temporal attention-augmented bilinear network for financial time-series data analysis,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 5, pp. 1407–1418, 2018.
 - [39] S. Y. Shih, F. K. Sun, and H. Lee, “Temporal pattern attention for multivariate time series forecasting,” *Machine Learning*, vol. 108, no. 8, pp. 1421–1441, 2019.
 - [40] J. Hu and W. Zheng, “Multistage attention network for multivariate time series prediction,” *Neurocomputing*, vol. 383, pp. 122–137, 2020.
 - [41] J. Bedi, “Attention based mechanism for load time series forecasting: AN-LSTM,” in *Proceedings of the 2020 International Conference on Artificial Neural Networks*, pp. 838–849, Bratislava, Slovakia, 2020.
 - [42] E. Chong, C. Han, and F. C. Park, “Deep learning networks for stock market analysis and prediction: methodology, data representations, and case studies,” *Expert Systems with Applications*, vol. 83, pp. 187–205, 2017.
 - [43] R. Singh and S. Srivastava, “Stock prediction using deep learning,” *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 18569–18584, 2017.

- [44] A. M. Ozbayoglu, M. U. Gudelek, and O. B. Sezer, "Deep learning for financial applications: a survey," *Applied Soft Computing*, vol. 93, Article ID 106384, 2020.
- [45] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-net: efficient channel attention for deep convolutional neural networks," in *Proceedings of the 2020 IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020.
- [46] R. K. Nayak, D. Mishra, and A. K. Rath, "A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices," *Applied Soft Computing*, vol. 35, pp. 670–680, 2015.
- [47] Y. Chen and Y. Hao, "A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction," *Expert Systems with Applications*, vol. 80, pp. 340–355, 2017.
- [48] W.-C. Chiang, D. Enke, T. Wu, and R. Wang, "An adaptive stock index trading decision support system," *Expert Systems with Applications*, vol. 59, pp. 195–207, 2016.
- [49] X.-d. Zhang, A. Li, and R. Pan, "Stock trend prediction based on a new status box method and AdaBoost probabilistic support vector machine," *Applied Soft Computing*, vol. 49, pp. 385–398, 2016.
- [50] I. Marković, M. Stojanović, J. Stanković et al., "Stock market trend prediction using AHP and weighted kernel LS-SVM," *Soft Computing*, vol. 21, no. 18, pp. 5387–5398, 2017.
- [51] L. Lei, "Wavelet neural network prediction method of stock price trend based on rough set attribute reduction," *Applied Soft Computing*, vol. 62, pp. 923–932, 2018.
- [52] A. Picasso, S. Merello, Y. Ma, L. Oneto, and E. Cambria, "Technical analysis and sentiment embeddings for market trend prediction," *Expert Systems with Applications*, vol. 135, pp. 60–70, 2019.
- [53] I. R. Parray, S. S. Khurana, M. Kumar, and A. A. Altalbe, "Time series data analysis of stock price movement using machine learning techniques," *Soft Computing*, vol. 24, no. 21, pp. 16509–16517, 2020.
- [54] Y. Xu, C. Yang, S. Peng, and Y. Nojima, "A hybrid two-stage financial stock forecasting algorithm based on clustering and ensemble learning," *Applied Intelligence*, vol. 50, no. 11, pp. 3852–3867, 2020.
- [55] R. K. Dash, T. N. Nguyen, K. Cengiz, and A. Sharma, "Fine-tuned support vector regression model for stock predictions," *Neural Computing and Applications*, pp. 1–15, 2021.
- [56] Q. Liang, W. Rong, J. Zhang, J. Liu, and Z. Xiong, "Restricted Boltzmann machine based stock market trend prediction," in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 1380–1387, Anchorage, AK, USA, 2017.
- [57] H. Y. Kim and C. H. Won, "Forecasting the volatility of stock price index: a hybrid model integrating LSTM with multiple GARCH-type models," *Expert Systems with Applications*, vol. 103, pp. 25–37, 2018.
- [58] B. Moews, J. M. Herrmann, and G. Ibikunle, "Lagged correlation-based deep learning for directional trend change prediction in financial time series," *Expert Systems with Applications*, vol. 120, pp. 197–206, 2019.
- [59] H. Li, J. Hua, J. Li, and G. Li, "Stock forecasting model FS-LSTM based on the 5G internet of things," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 7681209, 7 pages, 2020.
- [60] J. Zhao, D. Zeng, S. Liang, H. Kang, and Q. Liu, "Prediction model for stock price trend based on recurrent neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 745–753, 2021.
- [61] O. B. Sezer and A. M. Ozbayoglu, "Algorithmic financial trading with deep convolutional neural networks: time series to image conversion approach," *Applied Soft Computing*, vol. 70, pp. 525–538, 2018.
- [62] M. Wen, P. Li, L. Zhang, and Y. Chen, "Stock market trend prediction using high-order information of time series," *IEEE Access*, vol. 7, pp. 28299–28308, 2019.
- [63] S. Barra, S. M. Carta, A. Corriga, A. S. Podda, and D. R. Recupero, "Deep learning and time series-to-image encoding for financial forecasting," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 3, pp. 683–692, 2020.
- [64] J. Long, Z. Chen, W. He, T. Wu, and J. Ren, "An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market," *Applied Soft Computing*, vol. 91, Article ID 106205, 2020.
- [65] Y. Hao and Q. Gao, "Predicting the trend of stock market index using the hybrid neural network based on multiple time scale feature learning," *Applied Sciences*, vol. 10, no. 11, p. 3961, 2020.
- [66] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A CNN-LSTM-based model to forecast stock prices," *Complexity*, vol. 2020, Article ID 6622927, 10 pages, 2020.
- [67] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: a search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2016.
- [68] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 2011–2023, 2020.
- [69] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowledge and Information Systems*, vol. 60, no. 2, pp. 617–663, 2019.
- [70] Y. Yi, Y. Shi, W. Wang, G. Lei, J. Dai, and H. Zheng, "Combining boundary detector and SND-SVM for fast learning," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 3, pp. 689–698, 2021.
- [71] W. Zhou, S. Qiao, Y. Yi, N. Han, Y. Chen, and G. Lei, "Automatic optic disc detection using low-rank representation based semi-supervised extreme learning machine," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 1, pp. 55–69, 2020.
- [72] M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, and S. Wang, "Learning graph-based poi embedding for location-based recommendation," in *Proceedings of the 25th ACM International Conference on Information and Knowledge Management*, Indianapolis, IN, USA, 2016.

Research Article

Cluster Analysis and Visualization for the Legend of the Condor Heroes Based on Social Network

Chao Fan ^{1,2}, Zhihui Yang^{1,2} and Yuyi Yuan^{1,2}

¹The School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China

²Jiangsu Key Laboratory of Media Design and Software Technology, Jiangnan University, Wuxi 214122, China

Correspondence should be addressed to Chao Fan; fanchao@jiangnan.edu.cn

Received 24 May 2021; Revised 14 June 2021; Accepted 27 June 2021; Published 7 July 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Chao Fan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Legend of the Condor Heroes (LCH) is one of the fifteen well-known Wuxia novels penned by Jin Yong. It portrays a number of characters in the background of the Southern Song Dynasty. In this research, we attempt to analyze the relationship of characters in LCH based on social network, including network feature analysis, cluster analysis, and data visualization. Moreover, the approach can be extended to other literary works because our research provides a general framework for analyzing character relationships. We first perform lexical analysis on the corpus to extract character names and then utilize co-word analysis to build a social network of character relationships. We reckon characters as nodes and count the cooccurrences of characters as weights of links. By applying the social network analysis of created network, we can obtain network features of LCH. Furthermore, a hierarchical clustering algorithm is implemented to study the structure of LCH network. Both the dendrogram and Venn diagram are used for data visualization. An improved approach of visualizing the clustering results is advanced in order to display the group and hierarchical structure better. The final experimental results demonstrate that the proposed method shows a good effect.

1. Introduction

The Legend of the Condor Heroes, written by JIN Yong, is one of the most reputable novels painting the fantasy world of Wuxia. Based on real historical events, this novel tells the story of the Southern Song Dynasty. It describes the legend of Guo Jing, who is destined to shoulder the responsibility of the country and grow into a hero with the help of Huang Rong. The novel portrays about a hundred dramatic characters.

Relationships of characters can be explored from the perspective of social network. This paper tries to build a character relationship network based on the co-word analysis. It is well known that the cooccurrence of two characters' names in a paragraph or a sentence indicates a connection between them in some ways. Hence, the co-word analysis technology can be applied to the construction of character relationship network. At the beginning, textual analysis of LCH corpus is carried out by natural language processing such as segmentation, part-

of-speech tagging, etc. Utilizing the character names and their cooccurrences, we are able to build a weighted undirected network. Many network features are calculated to describe the characteristics of LCH network structure. Furthermore, the agglomerative hierarchical clustering method is implemented to analyze the relationships of characters. Some interesting conclusions of the novel can be drawn. Finally, an improved approach by employing the Venn diagram is proposed to visualize clustering results. By conducting cluster analysis and data visualization, not only does this work facilitate the exhibition of character relationships in LCH, but also it provides a framework for analyzing characters in the literary work.

The content of this paper can be presented in the following sections. Related studies are reviewed in Section 2. Social network construction is elaborated in Section 3. Section 4 states experimental results, discussions of cluster analysis, and data visualization. Section 5 gives the conclusion and future work.

2. Related Work

Quantitative research is an effective strategy to analyze literature, among which co-word analysis attracts the attention of many scholars. Co-word analysis is an important technology to build a social network. The co-word analysis approach was developed by French bibliographers [1] and soon introduced to other fields [2–4]. Some researchers [5] utilized the co-word analysis to study the theme of publication in the field of scientometrics. Their research results suggest that topics of publications often follow hot spots. Zhu and his team [1] concentrated on the scientific literature concerning social computing and found some hot research topics by using the co-word analysis. Gan et al. [6] leveraged co-word analysis tools to analyze the medical subject headings (MeSH) terms of epilepsy genetics. Nguyen [7] paid attention to the interdisciplinary literature on non-biomedical topics from 1987 to 2017. A large-scale co-word analysis was exploited for mapping knowledge. Another research [8] focused on the evolution of medical tourism. It analyzed the data from the Web of Science and Scopus based on co-word analysis and gave a visualization of all subfields. According to recent research [9], open data were employed to investigate knowledge areas, themes, and future research based on co-word analysis. Khaldi and Prado-Gascó [10] discussed international cooperation issues on migration by using co-word analysis of articles in Web of Science. They gave some frequent keywords and disclosed that most active authors in this field came from a few countries, such as the United States, the United Kingdom, Germany, etc.

Several authors explored the relationships of characters of literature by utilizing social network analysis, cluster analysis, and data visualization. Zhao et al. [11] extracted character relationships from Chinese literary based on social network. They discovered that the character network in Chinese literature had apparent small-world property and limited power-law distribution. Wang et al. [12] proposed a research method to analyze the historical characters in the Romance of the Three Kingdoms (RTK). They created a network according to the cooccurrence of character names and found some interesting phenomena in the RTK. Fan [13] established a character relationship network of the Dream of the Red Chamber. They calculated the network features and performed a cluster analysis. Some interesting conclusions can be drawn through the visualization of experimental results. Hu [14] made an effort to construct the character network in Records of the Three Kingdoms. Natural language processing (NLP) tools were taken into account to handle the text at first. Then segmentation and part-of-speech tagging were carried out. A custom dictionary was adopted to process the ancient Chinese text. Anaphora resolution was also used to determine which character a pronoun is referring to. Besides, the author applied a k-kernel decomposition approach to analyze the significant characters in the work. Xu [15] studied the classical literature Zuo Zhuan in the Pre-Qin period from the perspective of language network and social network. Both ancient and modern Chinese versions were selected as research objects. Language network considered the literature as language data

and preprocessed it by NLP technologies, whereas social network reckoned the work as historical documents and focused on the character relationships.

3. Creation of Social Network

3.1. Preprocessing of LCH Corpus. We downloaded a full text of the Legend of the Condor Heroes (LCH) in Chinese dataset (<https://www.uidzhx.com/Shtml604.html>). Data cleansing was done automatically and some punctuation errors were corrected manually. Also, mistakes of character names were revised according to the name list of LCH acquired from the Internet. After original dataset preprocessing, the created corpus can be used for lexical analysis.

The Chinese lexical analyzer ICTCLAS (<http://ictclas.nlpir.org/>) was employed to process the LCH corpus. Chinese segmentation was performed on each sentence and parts-of-speech were tagged for segmented words. Exploiting the results of segmentation and part-of-speech tagging, we succeeded in extracting the character names which were used for creating LCH network.

3.2. Construction of LCH Network. When generating nodes of the network, the full name, given name, and nickname of a character are regarded as one node. For instance, “Hong Qigong,” “Qigong,” and “Bei Gai” represent the same character. The cooccurrence of two names in a specified area (a paragraph or a sentence) denotes a link between two characters. The corresponding frequency indicates the weight of a link. In this paper, the contextual area is set as the paragraph.

We built a weighted undirected network to describe the relationships of characters in LCH based on co-word analysis. The established network incorporates 90 nodes and 1,013 links. A visualized LCH network is depicted in Figure 1 where degrees of nodes are ranked in the top 50. The size of a node represents the degree of the node.

3.3. Network Features

3.3.1. Degree Distribution. The degree of a node is defined as the count of its neighbors. The degree of a character means how many other characters are linked to him or her, thus implying the significance of a character. In the novel, a character may be directly connected to other characters or talked about in others’ conversations.

As can be seen from Figure 1, Guo Jing has the biggest value of the degree. The top ten people in degree ranking are as follows: Guo Jing, Huang Rong, Yang Kang, Qiu Chuji, Ke Zhene, Wanyan Honglie, Zhu Cong, Huang Yaoshi, Ouyang Feng, and Tuolei. We calculated the average degree of LCH network and obtained a value of 22.51. This number shows that a character in LCH is linked to an average of other 22.5 characters. That is to say, characters in LCH network are highly connected to each other.

Degree distribution $p(k)$ offers the probability that a node is related to k other nodes [16]. The degree distribution of LCH network is presented in Figure 2. A

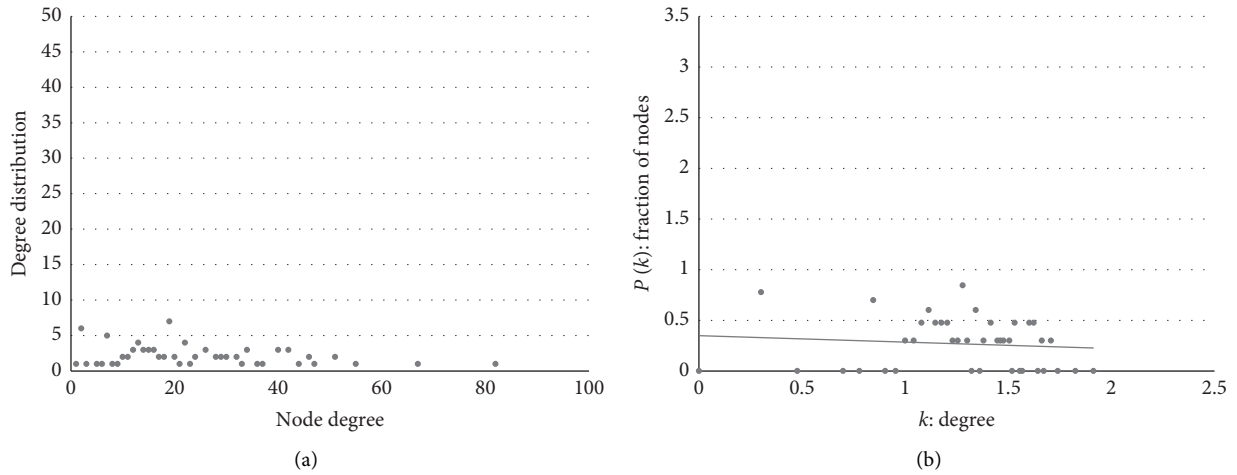


FIGURE 2: Degree distribution of LCH network and its log-log plot.

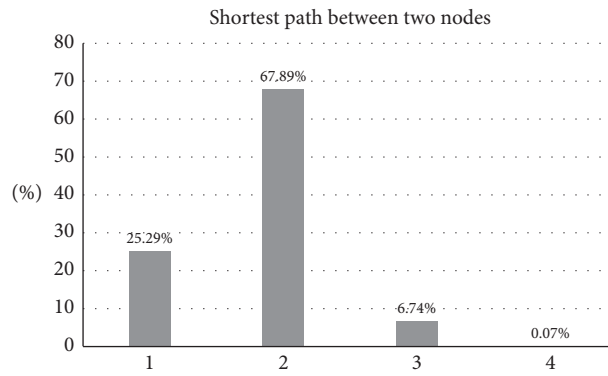


FIGURE 3: Distribution of shortest path length in LCH network.

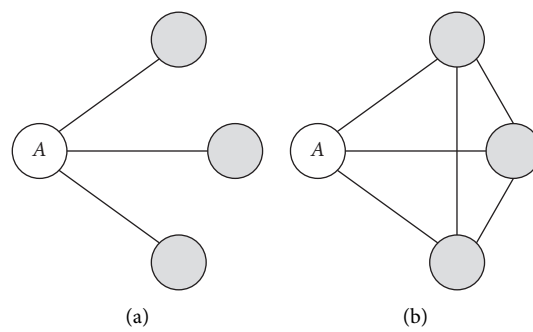


FIGURE 4: A comparison of clustering coefficient for node A.

TABLE 1: Network features of LCH network and random network.

	Number of nodes	Number of links	Clustering coefficient
LCH network	90	1013	0.7262
Random network	90	1013	0.2524

TABLE 2: Top 10 characters in centrality ranking.

Ranking	Degree centrality	Betweenness centrality	Closeness centrality
1	Guo Jing (0.9213)	Guo Jing (0.2323)	Guo Jing (0.9271)
2	Huang Rong (0.7528)	Huang Rong (0.0863)	Huang Rong (0.7946)
3	Yang Kang (0.6180)	Wanyan Honglie (0.0551)	Yang Kang (0.7177)
4	Qiu Chuji (0.5730)	Yang Kang (0.0484)	Qiu Chuji (0.7008)
5	Ke Zhene (0.5730)	Yang Tiexin (0.0333)	Ke Zhene (0.7008)
6	Wanyan Honglie (0.5281)	Ke Zhene (0.0296)	Wanyan Honglie (0.6794)
7	Zhu Cong (0.5169)	Qiu Chuji (0.0277)	Zhu Cong (0.6742)
8	Huang Yaoshi (0.5169)	Tuolei (0.0264)	Huang Yaoshi (0.6593)
9	Ouyang Feng (0.4944)	Tie Muzhen (0.0241)	Tuolei (0.6544)
10	Tuolei (0.4719)	Hong Qigong (0.0224)	Ouyang Feng (0.6496)

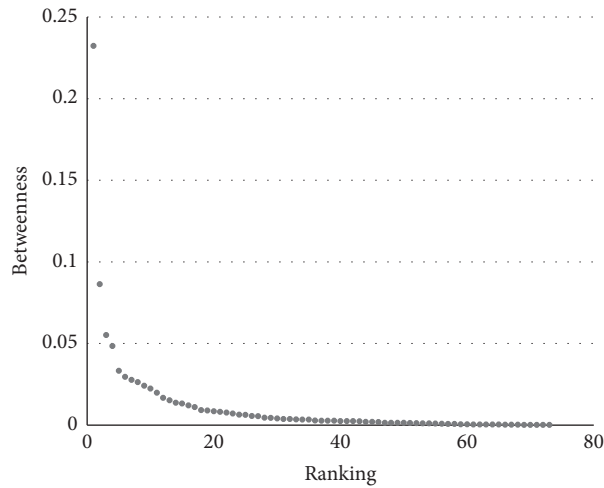


FIGURE 5: Distribution of betweenness for characters in LCH.

Closeness centrality [7] is specified by the reciprocal of the sum of the length of the shortest paths between the node and all other nodes in the network. It has the capability to estimate the degree of closeness for a node to reach other nodes in the network. The values of the top 10 closeness centralities in LCH network are specified in the fourth column of Table 2. Characters with high closeness centralities may rapidly communicate with others in the network.

According to Table 2, seven characters appear in the three columns of centrality ranking. They are Guo Jing, Huang Rong, Yang Kang, Qiu Chuji, Ke Zhene, Wanyan Honglie, and Tuolei, who are considered as the important characters in the novel.

4. Cluster Analysis and Visualization

4.1. Hierarchical Clustering

4.1.1. Cooccurrence Matrix and Similarity Matrix. Two matrices are raised to provide data for clustering. Given the frequency that two nodes appear in the same context, the cooccurrence matrix can be constructed based on co-word analysis. An example of cooccurrence matrix for six major characters in the LCH is illustrated in Table 3. The number on the diagonal indicates the count of character names appearing in the novel.

Since the weight of a link is affected by its adjacent nodes' degrees, we build the similarity matrix by calculating Ochiai coefficient [22]. Formula (4) gives the similarity of sets X and Y :

$$\text{sim}(X, Y) = \frac{n(X \cap Y)}{\sqrt{n(X) \times n(Y)}} \quad (4)$$

Here, $n(X)$ represents the frequency of X . $n(X \cap Y)$ is the cooccurrence number of X and Y . By applying the Ochiai coefficient, the cooccurrence matrix can be transformed into a correlation matrix, which is a similarity matrix. The result of the similarity matrix for the six main characters is presented in Table 4.

4.1.2. Clustering Algorithm. Hierarchical clustering is utilized to create a tree-like structure. It embraces two strategies: divisive [23] and agglomerative [24] clustering. We devised an agglomerative algorithm to cluster LCH characters into groups. Similarity matrix based on Ochiai coefficient is taken as input dataset of clustering. Agglomerative method is selected to implement the clustering algorithm. To start with, we regard each character as one group representing the cluster. At each time, two clusters that have the largest similarity are merged into a

TABLE 3: Cooccurrence matrix of six characters.

Cooccurrence	Guo Jing	Huang Rong	Yang Kang	Mu Nianci	Huang Yaoshi	Ouyang Feng
Guo Jing	1995	1197	145	50	242	361
Huang Rong	1197	1856	117	84	160	298
Yang Kang	145	117	424	104	20	38
Mu Nianci	50	84	104	251	7	6
Huang Yaoshi	242	160	20	7	586	157
Ouyang Feng	361	298	38	6	157	860

TABLE 4: Similarity matrix of six characters.

Cooccurrence	Guo Jing	Huang Rong	Yang Kang	Mu Nianci	Huang Yaoshi	Ouyang Feng
Guo Jing	1	0.622062	0.157657	0.070658	0.223818	0.275605
Huang Rong	0.622062	1	0.131891	0.12307	0.15342	0.235873
Yang Kang	0.157657	0.131891	1	0.318797	0.040123	0.062929
Mu Nianci	0.070658	0.12307	0.318797	1	0.018252	0.012914
Huang Yaoshi	0.223818	0.15342	0.040123	0.018252	1	0.221157
Ouyang Feng	0.275605	0.235873	0.062929	0.012914	0.221157	1

larger cluster. When all clusters are grouped into one cluster, the clustering process finishes. In addition, a cluster-stopping threshold can also be set so as to control the final number of clusters.

4.2. Result of Clustering. The clustering result of LCH characters can be visualized through a tree-like structure called dendrogram [21]. In this paper, we separated the whole dendrogram into groups with different numbers. A dendrogram example of LCH characters is depicted in Figure 6.

According to the figure, six groups are marked by red lines. Most characters in H1 are from “Quanzhen Taoism,” such as Ma Yu, Wang Chuyi, Liu Chuxuan, etc. H2 contains the most important characters of the novel, like Guo Jing, Huang Rong, Yang Kang, and Mu Nianci. Many martial arts masters are also included in this group: Huang Yaoshi, Ouyang Feng, Hong Qigong, and Yideng. H3 incorporates the “seven heroes” and the “Twice Foul Dark Wind” in the novel. Both of them are closely linked groups and can merge into a bigger group. Further, most characters in H4 are from a gang named “Gai.” They are homeless but organized by some leaders. Characters in the H5 group are from the Mongolian nation, e.g., Tie Muzhen (Genghis Khan) and Tuolei (Tolui). Finally, H6 is composed of “Four ghosts of the Yellow River” and the remainder. Six parts make up the Wuxia world of LCH.

4.3. Improved Visualization Approach. The result of hierarchical clustering is in essence a binary tree. As shown in Figure 7, a tree can be displayed in the form of a Venn diagram [25] with the hierarchical structure. The nodes C, D, and E are leaf nodes, representing characters in the novel, whereas the nodes A and B refer to groups with

different sizes. Using the Venn diagram to display hierarchical clustering results has many advantages. For instance, it can clearly express the group features and hierarchical structure.

Nonetheless, there may be some disadvantages to this visualization approach. Since 90 characters appear in the novel, the number of nodes should be 90 and the number of groups would be 89. As a result, there will be 179 circles in the Venn diagram, leading to a bad visualization effect (see left side in Figure 8).

When displaying the characters, it is not necessary to draw all group circles in the hierarchy. In order to reduce the number of group circles, a merge operation is introduced in this research, which is shown in Figure 8. By removing the group circle B and G, or even A and F, two forms of diagrams are obtained in the right side of Figure 8.

In hierarchical clustering, we can control the number of clusters and merge nodes with the same cluster label. A visualization of LCH characters in five clusters is given in Figure 9. Different groups are filled with different colors. According to the figure, our proposed visualization approach is capable of showing the hierarchical structure of clusters in a clear manner. From the visualized clustering results, we can discover two major groups: Chinese groups and Mongolian groups. As Guo Jing grew up in Mongolia, the novel portrays a number of characters in Mongolian groups in the early chapters. In Chinese groups, the largest group contains the most significant roles in the LCH. Then, “the seven heroes” group and the largest group can merge into a bigger one. As the merge process goes on, the “Quanzhen Taoism” and “Four ghosts of the Yellow River” groups also joined the Chinese groups. Finally, characters in both Chinese and Mongolian groups make up the Wuxia world of LCH.

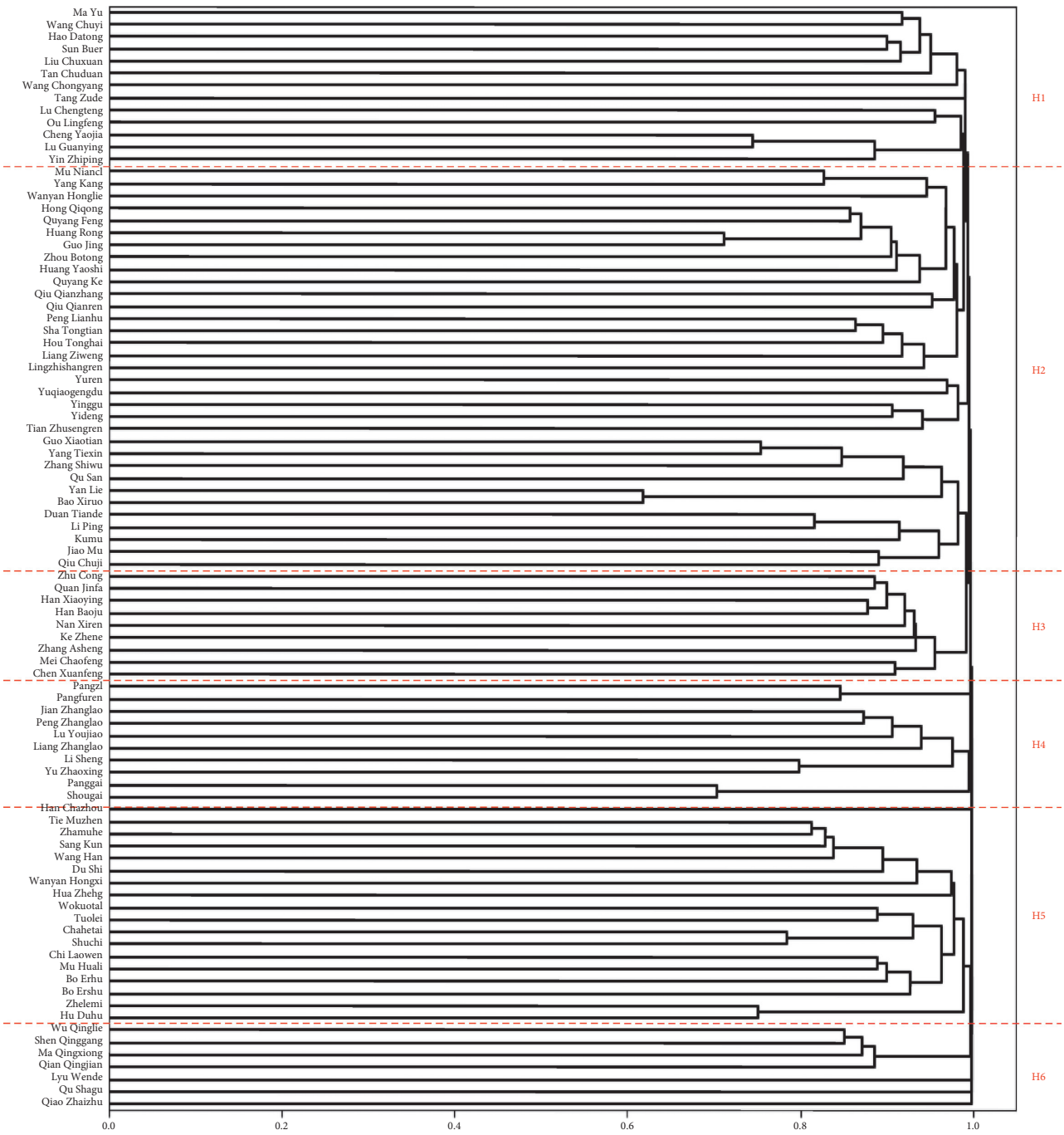


FIGURE 6: Dendrogram of clustering result for LCH network.

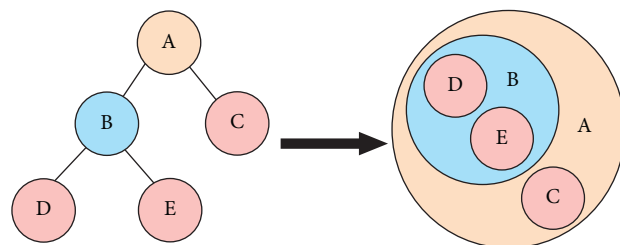


FIGURE 7: Two visualization approaches of tree structure.

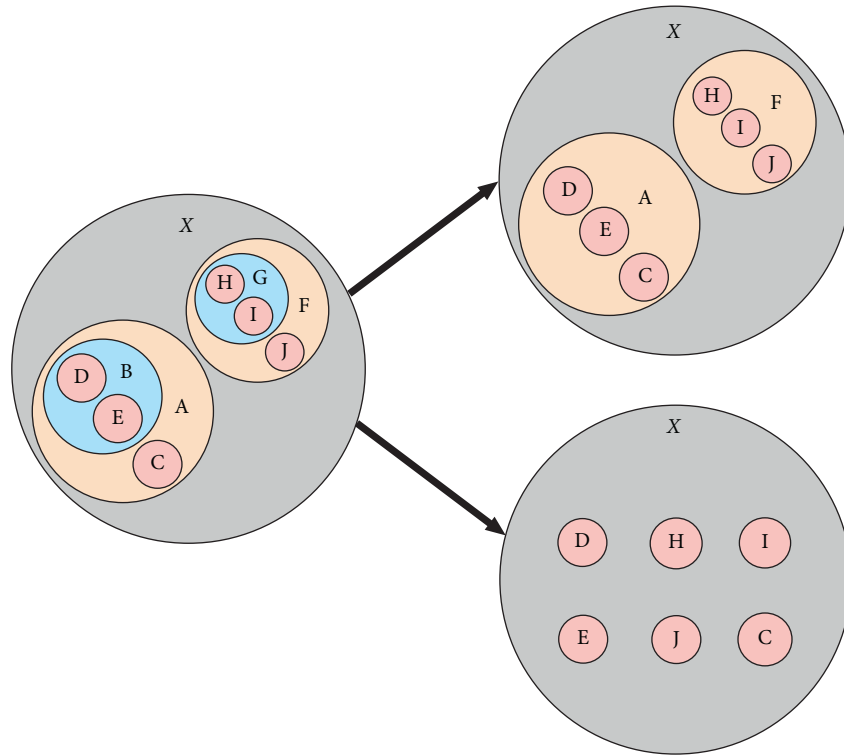


FIGURE 8: The merge operation of small groups in the Venn diagram.

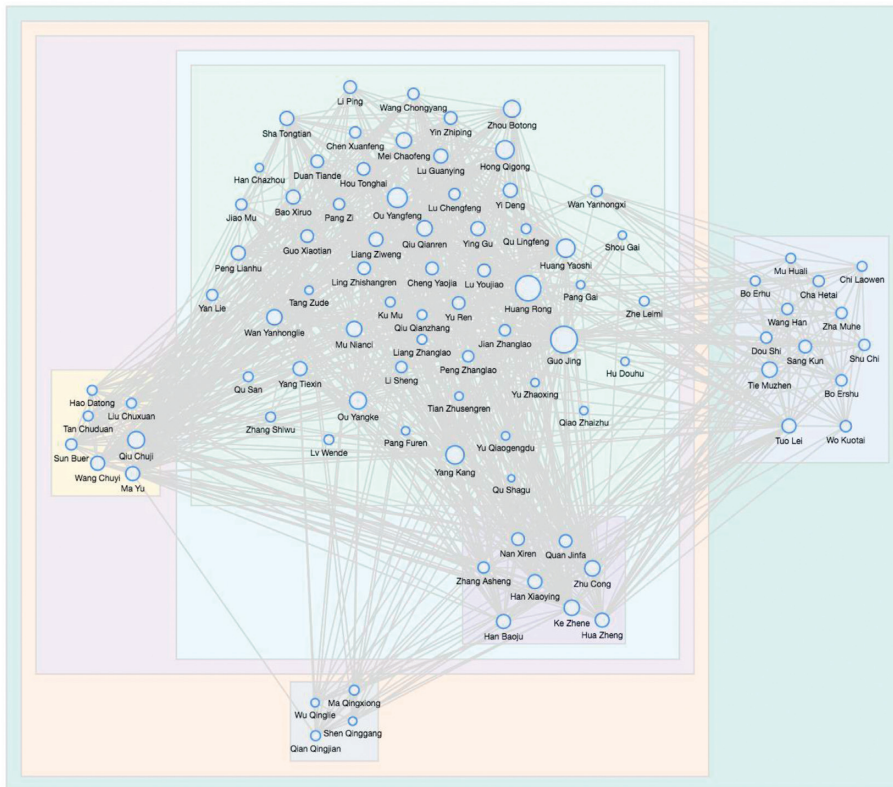


FIGURE 9: An example of the improved visualization method.

5. Conclusions

This paper concentrates on the analysis and visualization of character relationships in the “Legend of the Condor Heroes (LCH).” The Chinese version of LCH was selected as the research object. At the beginning, Chinese segmentation and part-of-speech tagging were enforced to preprocess the LCH corpus. Adopting the technology of co-word analysis, character names and cooccurrence of names in the context were thought to be nodes and links in the network. We then constructed a weighted undirected network of relationships of characters in LCH. Based on the social network analysis, computation of network features was completed such as centrality, clustering coefficient, density, etc. Furthermore, hierarchical clustering was performed and a dendrogram was sketched. An improved visualization approach of Venn diagram was also proposed to exhibit the effect of hierarchical clustering. From the experimental result, we can identify two major groups (Chinese and Mongolian groups) and the hierarchical structure in Chinese groups.

However, there are some disadvantages in this research. The cooccurrence of character names in the context may not represent the real relationships of characters. Nevertheless, the proposed quantitative research method provides a new perspective to analyze the characters in the novel. In the future, coreference resolution will be taken into consideration to build accurate relationships of characters. Moreover, we expect to explore the semantic analysis of texts to refine the experimental result.

Data Availability

The LCH dataset used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Youth Foundation of Basic Science Research Program of Jiangnan University, 2019 (no. JUSRP11962), and the High-Level Innovation and Entrepreneurship Talents Introduction Program of Jiangsu Province of China, 2019.

References

- [1] Q. Zhu, X. Peng, and X. Liu, “Research topics in social computing area based on co-word analysis,” *Information Studies: Theory & Application*, vol. 12, pp. 7–11, 2012.
- [2] M. Callon, J.-P. Courtial, W. A. Turner, and S. Bauin, “From translations to problematic networks: an introduction to co-word analysis,” *Social Science Information*, vol. 22, no. 2, pp. 191–235, 1983.
- [3] Q. He, “Knowledge discovery through co-word analysis,” *Library Trends*, vol. 48, no. 1, pp. 133–159, 1999.
- [4] X. Wang, “Formation and evolution of science knowledge network (I): a new research method based on co-word network,” *Journal of the China Society for Scientific and Technical Information*, vol. 28, no. 4, pp. 599–605, 2009.
- [5] S. Ravikumar, A. Agrahari, and S. N. Singh, “Mapping the intellectual structure of scientometrics: a co-word analysis of the journal scientometrics (2005–2010),” *Scientometrics*, vol. 102, no. 1, pp. 929–955, 2015.
- [6] J. Gan, Q. Cai, P. Galer et al., “Mapping the knowledge structure and trends of epilepsy genetics over the past decade: a co-word analysis based on medical subject headings terms,” *Medicine*, vol. 98, no. 32, Article ID e16782, 2019.
- [7] D. Nguyen, “Mapping knowledge domains of non-biomedical modalities: a large-scale co-word analysis of literature 1987–2017,” *Social Science & Medicine*, vol. 233, pp. 1–12, 2019.
- [8] A. de la Hoz-Correa, F. Muñoz-Leiva, and M. Bakucz, “Past themes and future trends in medical tourism research: a co-word analysis,” *Tourism Management*, vol. 65, pp. 200–211, 2018.
- [9] D. Corrales-Garay, M. Ortiz-de-Urbina-Criado, and E.-M. Mora-Valentín, “Knowledge areas, themes and future research on open data: a co-word analysis,” *Government Information Quarterly*, vol. 36, no. 1, pp. 77–87, 2019.
- [10] H. Khaldi and V. Prado-Gascó, “Bibliometric maps and co-word analysis of the literature on international cooperation on migration,” *Quality and Quantity*, vol. 2021, no. 1, pp. 1–25, 2021.
- [11] J. Zhao, L. Zhang, Q. Zhu, and G. Zhou, “Extracting and analyzing social networks from Chinese literary,” *Journal of Chinese Information Processing*, vol. 31, no. 2, pp. 99–106, 2017.
- [12] Y. Wang, J. Yu, and C. Zhao, “Research on application of co-word analysis on relationships of characters in the romance of the three kingdoms,” *Information Research*, vol. 7, pp. 52–56, 2017.
- [13] C. Fan, “Research on relationships of characters in the dream of the red chamber based on co-word analysis,” *ICIC Express Letters Part B: Applications*, vol. 11, no. 5, pp. 1–8, 2020.
- [14] Y. Hu, *Construction and Analysis of Social Networks in Chinese Literature: A Case Study on Three Kingdoms*, Zhejiang Gongshang University, Hangzhou, China, 2019.
- [15] C. Xu, *Research into the Language Network and Social Network of Zuo Zhuan*, Nanjing Normal University, Nanjing, China, 2014.
- [16] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [17] C. Fan and Y. Li, “Coward and cluster analysis for the romance of the three kingdoms,” *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5553635, 8 pages, 2021.
- [18] D. J. Watts and S. H. Strogatz, “Collective dynamics of “small-world” networks,” *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [19] B. Bollobás, *Random Graphs*, Cambridge University Press, Cambridge, UK, 2nd edition, 2001.
- [20] M. E. J. Newman, *Networks*, Oxford University Press, Oxford, UK, 2nd edition, 2018.
- [21] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge, UK, 1994.
- [22] Q. Zhou and L. Leydesdorff, “The normalization of occurrence and co-occurrence matrices in bibliometrics using cosine similarities and Ochiai coefficients,” *Journal of the Association for Information Science & Technology*, vol. 67, no. 11, pp. 1–25, 2016.

- [23] A. Ishizaka, B. Lokman, and M. Tasiou, “A stochastic multi-criteria divisive hierarchical clustering algorithm,” *Omega*, vol. 103, Article ID 102370, 2021.
- [24] N. Liu, Z. Xu, X.-J. Zeng, and P. Ren, “An agglomerative hierarchical clustering algorithm for linear ordinal rankings,” *Information Sciences*, vol. 557, pp. 170–193, 2021.
- [25] N. B. J. Bipin and M. S. Sarath, “Evolutionary clustering annotation of ortho-paralogous gene in a multi species using Venn diagram visualization,” *International Journal of Engineering & Technology*, vol. 7, no. 1, pp. 162–166, 2018.

Research Article

Human-Machine Cooperation and Path Planning for Complex Road Conditions

Guanghong Zhou 

Department of Control Technology, Wuxi Institute of Technology, Wuxi, Jiangsu 214121, China

Correspondence should be addressed to Guanghong Zhou; zhough@wxit.edu.cn

Received 19 May 2021; Revised 17 June 2021; Accepted 28 June 2021; Published 6 July 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Guanghong Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of the information age, the development of industrial robots is also advancing by leaps and bounds. In the scenes of automobile, medicine, aerospace, and public service, we have fully enjoyed the convenience brought by industrial robots. However, with the continuous development of industrial robot-related concepts and technologies, human-computer interaction and cooperation have become the development trend of industrial robot. In this paper, the human-machine cooperation and path optimization of industrial robot in a complex road environment are studied and analyzed. At the theoretical modeling level, firstly, the industrial robot is modeled and obstacle avoidance is analyzed based on the kinematics of industrial robot; thus, an efficient and concise collision detection model of industrial robot is proposed. At the algorithm level, in view of the complex road conditions faced by industrial robots, this paper will study and analyze the obstacle avoidance strategy of human-computer cooperation and real-time path optimization algorithm of industrial robots. Based on the virtual target point algorithm, this paper further improves the problem that the goal of the traditional path planning algorithm cannot be fully covered, so as to propose the corresponding improved path planning algorithm of industrial robots. In the experimental part, based on the existing industrial robot system, the human-machine cooperation and path planning system proposed in this paper are designed. The experimental results show that the algorithm proposed in this paper improves the accuracy of obstacle avoidance by about 10 points and the corresponding convergence speed by about 5% compared with the traditional algorithm and the experimental effect is remarkable.

1. Introduction

With the rapid development of industrial robots, more and more kinds of industrial robots come into people's daily life. Industrial robots can completely replace human beings to complete highly repetitive work. At the same time, they also further liberate productivity and improve production efficiency. However, with the continuous advancement of the concept of industrial robot and related technologies, the technology of industrial robot is developing in the direction of man-machine cooperation. The new man-machine cooperation mode will continue to break the boundaries between traditional industrial robots and workers, leading to the development trend of industrial robot technology [1–3]. The industrial robot based on human-computer cooperation makes it possible to cooperate with human beings without

absolute isolation by safety fence, thus further reducing the distance between human and machine, reducing the area required by industrial production, fully combining the advantages of human and machine, and realizing complementary advantages. In the actual production activities, let the robot complete the mechanical work with high repeatability, and let the human complete the work with relatively strong flexibility, which needs continuous optimization by manual work [4, 5]. However, compared with the traditional industrial robots, the development of human-computer cooperation has brought about the security problems of the system. When human beings cooperate with machines and industrial robots move in the production place, the collision avoidance between human and machines, between machines and machines, and between machines and other objects can hardly be guaranteed. Industrial robot

human-computer cooperation and corresponding path optimization algorithm has become an important research direction in the development of human-computer cooperation, and it is also the key difficulty to further promote this technology [6–8].

At present, the research on human-computer cooperation and path planning of industrial robot mainly focuses on the sensor level, which installs a large number of sensors on human and industrial robot, so as to realize mutual perception between industrial robot and human [9]. Sensor sensing of the current mainstream technology mainly used sensors to collect uncertain or unknown factors in the environment to capture human action. Through full analysis of the captured information, it generates the motion simulation representation of human and industrial robot and estimates human behavior by this simulation representation. Then, the actual distance between human and industrial robot is further judged, and the distance is summarized into the collision strategy of human and robot. At this time, when the corresponding collision risk is confirmed by industrial robot, the corresponding collision strategy will be activated, and the corresponding path trajectory of industrial robot will be generated, so as to realize the obstacle avoidance processing of human or related objects and machines [10–12]. In addition to the above-mentioned contact sensor technology, the current mainstream human-computer cooperation and path planning technology also includes contact motion capture system, but such systems often need to bear high operating costs when applied in actual scenes, and their customization degree is also high, which is not conducive to industrial promotion and use [13, 14]. Based on this, this paper will study and analyze the human-machine cooperation and path optimization of industrial robot in a complex road environment. Firstly, the industrial robot is modeled and obstacle avoidance is analyzed based on the kinematics of the industrial robot, so as to put forward an efficient and concise human-machine collision detection model of industrial robot. In view of the complex road conditions faced by industrial robots, this paper will study and analyze the obstacle avoidance strategy of human-computer cooperation and real-time path optimization algorithm of industrial robots. Based on the virtual target point algorithm, this paper further improves the problem that the goal of traditional path planning algorithm cannot be fully covered and puts forward the corresponding improved algorithm of industrial robot path planning [15]. In the experimental part, based on the existing industrial robot system, this paper designs the human-computer cooperation and path planning system. The experimental results show that the algorithm proposed in this paper can achieve the overall obstacle avoidance of industrial robot more efficiently than the traditional algorithm, and the planned path is more safe and efficient, which verifies the feasibility and superiority of the algorithm.

The structure of this paper is as follows: in the second section of this paper, the research status of human-computer cooperation and path planning algorithm of industrial robot is analyzed. The third section of this paper will focus on the optimization algorithm of human-computer cooperation

and path planning of industrial robot, aiming at the research and analysis of human-computer cooperation and path optimization of industrial robot in a complex road environment. In the fourth section of this paper, the simulation system is designed based on the algorithm proposed in this paper, the simulation experiments are carried out, and the experimental results are analyzed. The summary is presented in the final section.

2. Related Research: Research Status of Human-Computer Cooperation and Path Planning Algorithm of Industrial Robot

At present, the essence of the research on human-computer cooperation and path planning algorithm of industrial robot is the safety of human-computer cooperation. Based on this, a large number of researchers and R&D institutions have carried out research and analysis on it. At the physical level, relevant researchers in the United States and Europe have proposed to use physical materials with absorption elasticity to develop industrial robots, so as to reduce the damage caused by the collision between robots and human bodies. The robot combines the nonlinear control theory with the electric drive system to realize the dynamic control of the industrial robot, so as to further ensure the safety of workers. With the continuous development of this kind of robot technology, the new robot has the characteristics of smooth structure, small size, and lightweight. However, this kind of robot does not completely solve the safety problem of human-computer cooperation, it only reduces the corresponding damage, and its corresponding production cost and maintenance cost are relatively high [16–18]; relevant scholars have studied and analyzed the control algorithm, and the main research scope is focused on obstacle avoidance strategy and path planning. Asian scholars have proposed a motion planning algorithm for industrial robots based on risk index minimization, which divides the robot motion path and region into safe path planning, online real-time safe trajectory planning, and real-time obstacle avoidance control [19, 20]; American scholars have designed a distributed distance sensor, which combines the risk assessment algorithm of the sensor with the control of industrial robot and then allows the robot to make corresponding obstacle avoidance actions, so as to improve the safety of the system; European scientists propose a motion control strategy to ensure safety, which mainly uses camera technology to monitor the working area of robot, calculates the distance between industrial robot and human in real time, and then defines various behaviors of robot based on this distance to ensure production efficiency and human safety [21, 22]; relevant researchers have focused on the application of video monitoring technology. Asian scientists used the somatosensory camera technology to track the position of workers, so as to obtain the three-dimensional coordinates of human joints. At the same time, they obtained the accurate position information of the target object through image processing technology and finally achieved a safe man-machine cooperation relationship [23].

Based on the somatosensory camera technology, Japanese scholars proposed further optimization of tracking technology. They developed an algorithm to estimate the distance between the omnipoint and the obstacle in Cartesian space and used the truncated cone generated by the pixels in the depth image to estimate the distance, so as to quickly calculate the distance between the industrial robot and human beings, so as to improve the running speed of the whole system [24, 25].

3. Optimization Algorithm Analysis of Human-Machine Cooperation and Path Planning for Industrial Robot

This section will mainly discuss and analyze the algorithm optimization of the current industrial robot human-machine cooperation and path planning level, which mainly includes the optimization of two core algorithms; the corresponding are the industrial robot kinematics optimization modeling analysis and the industrial robot human-machine cooperation security strategy and path planning research. The corresponding system framework diagram is shown in Figure 1. From the diagram, we can see that the ultimate goal of human-computer cooperation of industrial robot is to protect the personal safety of the collaborators. The corresponding diagram contains the algorithm hardware and software components of the system, as well as the security strategy of the whole system.

3.1. Kinematics Optimization Modeling Analysis of Industrial Robot. In order to better solve the problem of human-computer cooperation and path planning of industrial robot, this section analyzes the kinematics modeling of industrial robot. In this paper, we focus on the axis robot as an example for analysis and research, its specific needs to go through the industrial robot position description, attitude description, and link description. In this paper, the collision strategy between industrial robot and human is considered in the actual modeling, and based on this, the kinematics optimization of industrial robot is carried out. The principle framework of the corresponding industrial robot kinematics model is shown in Figure 2.

In the corresponding position description part of the industrial robot, a certain space coordinate system is established based on three-dimensional space, and the points on the rigid body of the industrial robot are represented by three-dimensional Cartesian vector. In the corresponding orientation description level, the attitude representation is mainly based on the relative coordinates of the coordinate system. Assuming that $\{Q\}$ is corresponding to the original coordinate system, then the corresponding rigid body coordinate system corresponds to $\{P\}$, and the attitude of the rigid body can be expressed as the position of the coordinate system fixedly connected to the rigid body relative to the original coordinate system. In the description of the pose of the corresponding industrial robot, this paper focuses on its corresponding position and pose. At this time, the corresponding rigid body feature points such as the

centroid or the center point are usually taken as the origin of the rigid body coordinate system. Compared with the original coordinate system $\{Q\}$, the corresponding rigid body coordinate system can be expressed as the following matrix 1, and the corresponding matrix has been homogeneous:

$$Q_A^B = \begin{bmatrix} p_i & o_i & p_i & o_i \\ p_j & o_j & p_j & o_j \\ p_z & o_z & p_z & o_z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (1)$$

In the corresponding linkage description, it is mainly subject to the structure of industrial robot. The corresponding structure of conventional industrial robot is spatial open-chain structure. The corresponding connecting rod is processed in series through some column joints. The first link of corresponding industrial robot is recognized as the base of industrial robot by default, and the latter link of the corresponding group is regarded as the end actuator. The corresponding modeling rules of connecting rod parameters based on this rule are as follows:

- (1) The corresponding z -axis direction in the corresponding linkage system coordinate system is set as the axis of the joint of industrial robot.
- (2) The corresponding origin is assumed to be the vertical line corresponding to the z -axis direction.
- (3) The corresponding industrial robot x -axis is set as the z -axis, and the vertical line of its adjacent z -axis and the corresponding direction is determined by z -axis.
- (4) The setting of the corresponding y -axis follows the right-hand rule and is determined by the right-hand rule.

Based on the preparation of the above-mentioned related modeling, the collision strategy between industrial robot and human is considered and the kinematics model of industrial robot is optimized. In the actual modeling, the obstacles in the environment are regarded as operators in the collaborative environment, corresponding to the real-time updated obstacles. Aiming at the attitude and position information model of the industrial robot and human body established before, the precise three-dimensional model of industrial robot and human can be obtained. Based on the above model, the collision detection model is built based on the bounding box collision algorithm. The corresponding modeling methods are as follows:

- (1) The collision detection of irregular obstacles in the environment is transformed into collision detection between corresponding regular geometry and can be calculated quickly.
- (2) According to different types of bounding box algorithms, the corresponding envelope form is set as the AABB model.
- (3) When the corresponding industrial robot envelope is collected, the images of multiple perspectives are

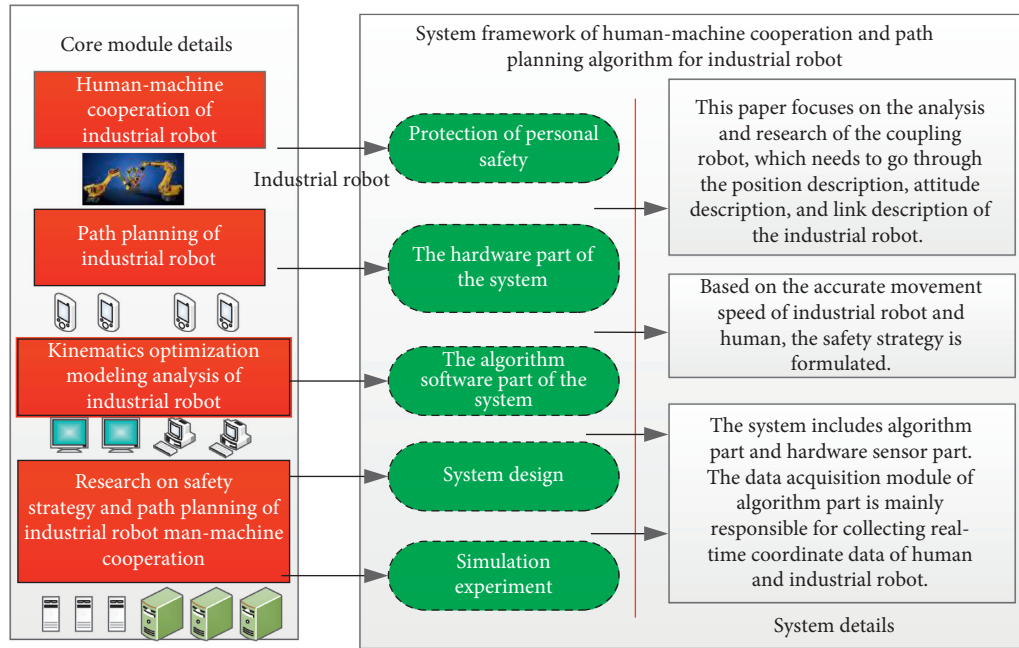


FIGURE 1: Framework of human-machine cooperation and path planning algorithm system for industrial robot.

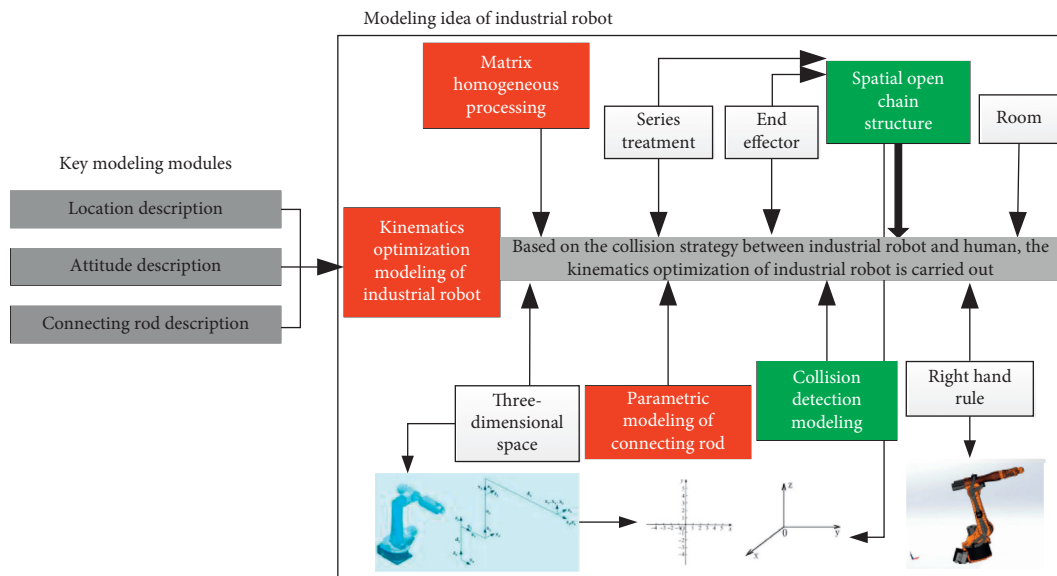


FIGURE 2: Schematic diagram of kinematic model of industrial robot.

collected, and the background is segmented and filtered, so as to get the segmented images of the robot from different perspectives.

- (4) The corresponding space segmentation images are subdivided into several basic elements, and the position information of the finite cylinder and the corresponding key nodes of industrial robot are determined according to the known relevant information.
- (5) A complete operational model of industrial robot is built.

The flowchart of the corresponding kinematic model of industrial robot is shown in Figure 3.

3.2. Analysis and Research on Safety Strategy and Path Planning of Industrial Robot Man-Machine Cooperation. Based on the above industrial robot kinematics modeling, considering the core algorithm of the whole system, namely, the safety strategy algorithm, this section will formulate the safety strategy based on the accurate movement speed of industrial robot and human. In the actual development of the corresponding safety strategy, the corresponding safety distance is set to 100 cm. When the distance between the corresponding industrial robot and human exceeds the safety distance of 100 cm, it indicates that the corresponding human is outside the working space of the industrial robot, and the corresponding industrial robot will work at a certain

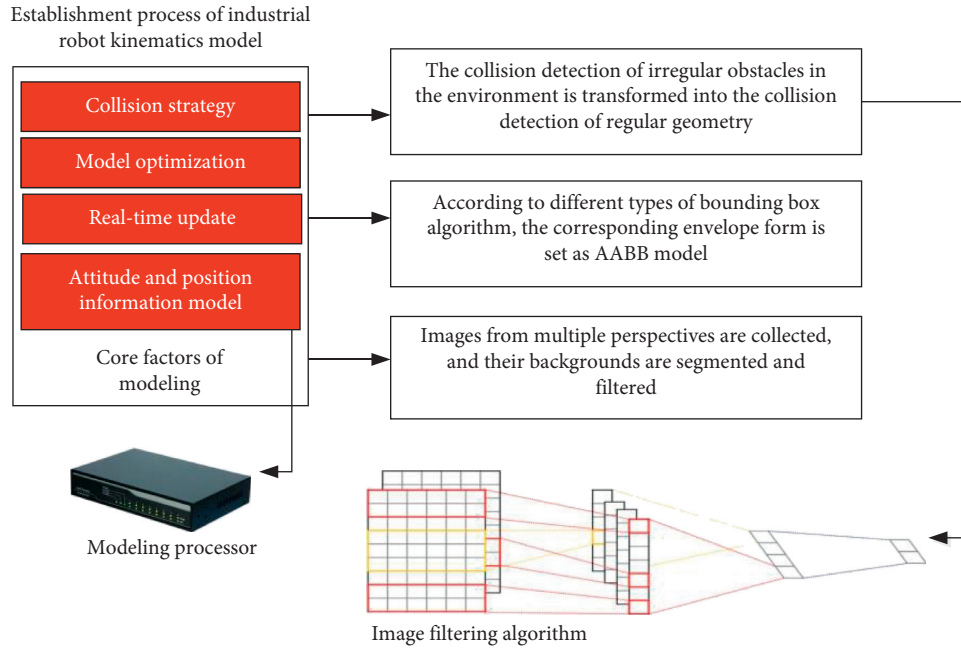


FIGURE 3: Flowchart of industrial robot kinematics model establishment.

speed v . When the safety distance between the corresponding industrial robot and human is less than 100 cm, but this kind of data only has a single digit, then the strategy thinks that the data at this time is the noise data of the sensor, and the industrial robot still works at a certain speed v . When the corresponding warning data shows periodic stability, it is considered that the distance between the industrial robot and human is too close, and the working speed of the industrial robot is reduced. When the safety distance between the industrial robot and human is shrinking, the industrial robot stops working. The corresponding real-time adjustment function of the industrial robot is shown in formula (2). In the formula, a represents the reduction percentage of the speed of the industrial robot, d represents the safe distance between human and industrial robot, d_1 and d_2 represent the distance between human and machine when the industrial robot begins to reduce the working speed and stops working, and V_1 and V_2 represent the deceleration function of the industrial robot. To a certain extent, it controls the speed of industrial robots.

$$a(d) = \begin{bmatrix} 0 & d > d_{\max} & 0 \\ \frac{(1 - v_1(d - 15))}{y} & d_{\min} < d < d_{\max} & v_1 \\ 1 & d < d_{\min} & v \end{bmatrix}. \quad (2)$$

The most important of the above-mentioned industrial robot safety strategies is the precise control of the speed of industrial robots. In this paper, the auxiliary functions V_1 and V_2 are used to control the speed of industrial robots in practice. This can improve the security of the whole system and reduce the mechanical wear of the industrial robot. The

corresponding speed precision control curve used in this paper is shown in Figure 4. From the figure, it can be seen that the safety of the system can be maximized by considering the corresponding speed, inertia, sharpness, and other attributes of industrial robots.

In the corresponding path planning algorithm level, we mainly consider improving the traditional form of artificial potential field. In this paper, the potential function is used to replace the traditional potential field force method for path planning of industrial robot, so as to achieve obstacle avoidance.

Through observation, analysis, and summary, the smaller the rotation angle of each joint of industrial robot is, the lower the probability of collision with human will be. Therefore, the improved gravitational potential function is shown in formula (3), where k represents the coefficient of joint angle distance, q represents the angle of each joint of industrial robot, and p represents the position of the end point of industrial robot.

$$M(p) = \left(\frac{k}{2}\right) * (\{p - p_i\}^2) + \left(\frac{k}{2}\right)k * \sum_{i=1}^n \{p - p_i\}. \quad (3)$$

Aiming at the problem that the traditional path planning algorithm can not reach the target, this paper introduces the distance between the end point and the target point of the industrial robot on the basis of the gravitational potential function. When the corresponding industrial robot is close to the corresponding target point, the distance between the end point and the target point is shrinking, which plays a drag role on the corresponding repulsive potential field. Thus, the problem of unreachable target is solved. Based on this, the flowchart of the corresponding path planning algorithm is shown in Figure 5. From Figure 5, we can see that the corresponding details are as follows:

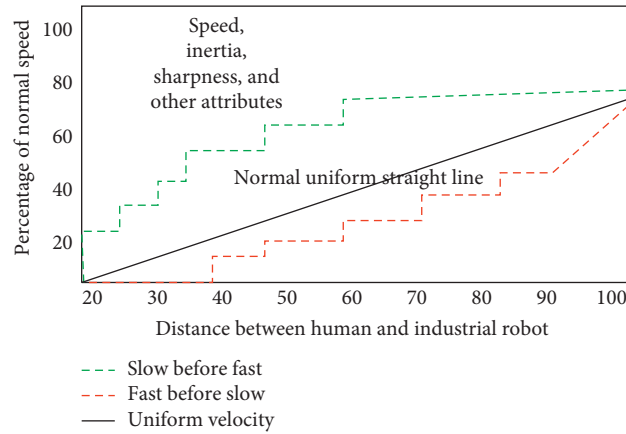


FIGURE 4: Speed precision control curve of industrial robot.

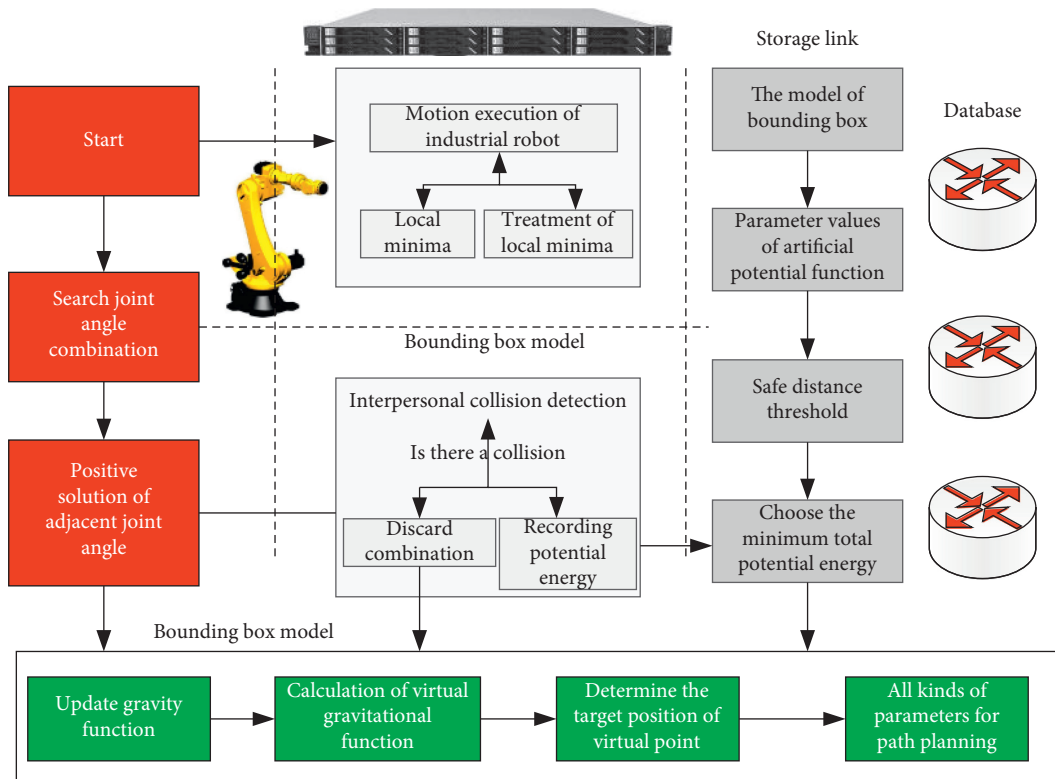


FIGURE 5: Flowchart of path planning algorithm for human-machine cooperation of industrial robot.

- (1) Get the initial joint angle of the industrial robot, establish the bounding box model through the corresponding sensor, and determine the corresponding parameters of the improved artificial potential function, the repulsion influence distance, and the corresponding safety distance threshold.
- (2) Combine the corresponding joints.
- (3) Using the forward kinematics equation of the industrial robot, the coordinates of all joints and end points of the industrial robot under each joint combination are obtained
- (4) Build the bounding box model and test the collision with the human bounding box.
- (5) Substituting the recording distance corresponding to step 4 into the function formula of gravity and repulsion, the potential energy corresponding to each combination of industrial robot is calculated and compared, and the joint angle combined with the smallest total potential energy is selected.
- (6) Path planning based on the above parameters.

4. System Design and Simulation Experiment

In order to verify the superiority of the algorithm, the overall design of the system is carried out first, and the corresponding overall design architecture is shown in Figure 6. It

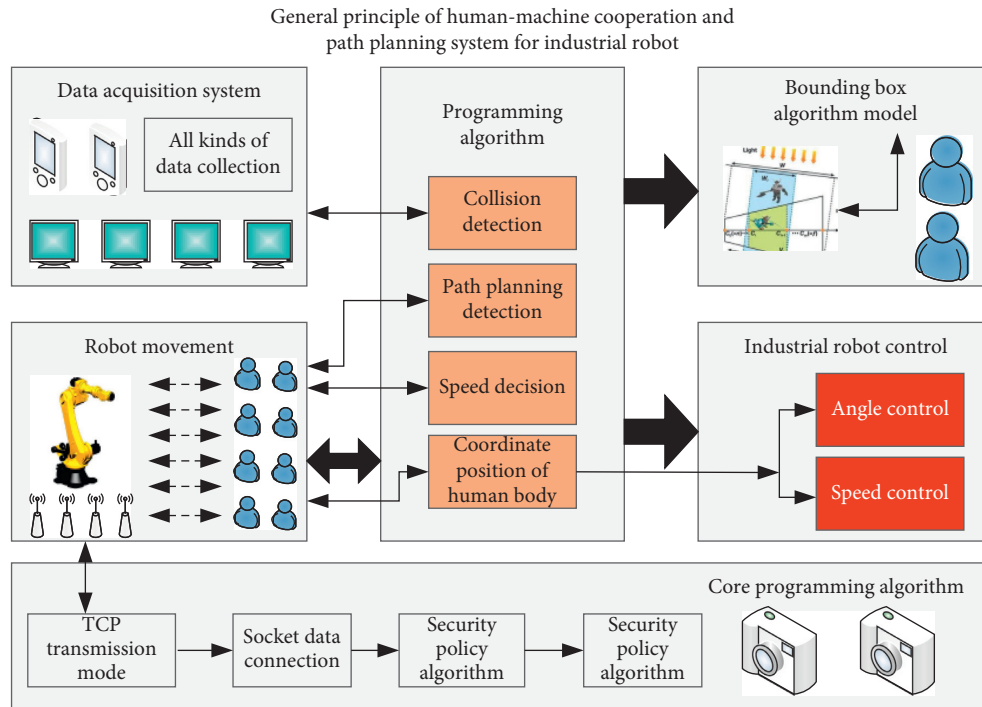


FIGURE 6: General principle framework of industrial robot human-machine cooperation and path planning system.

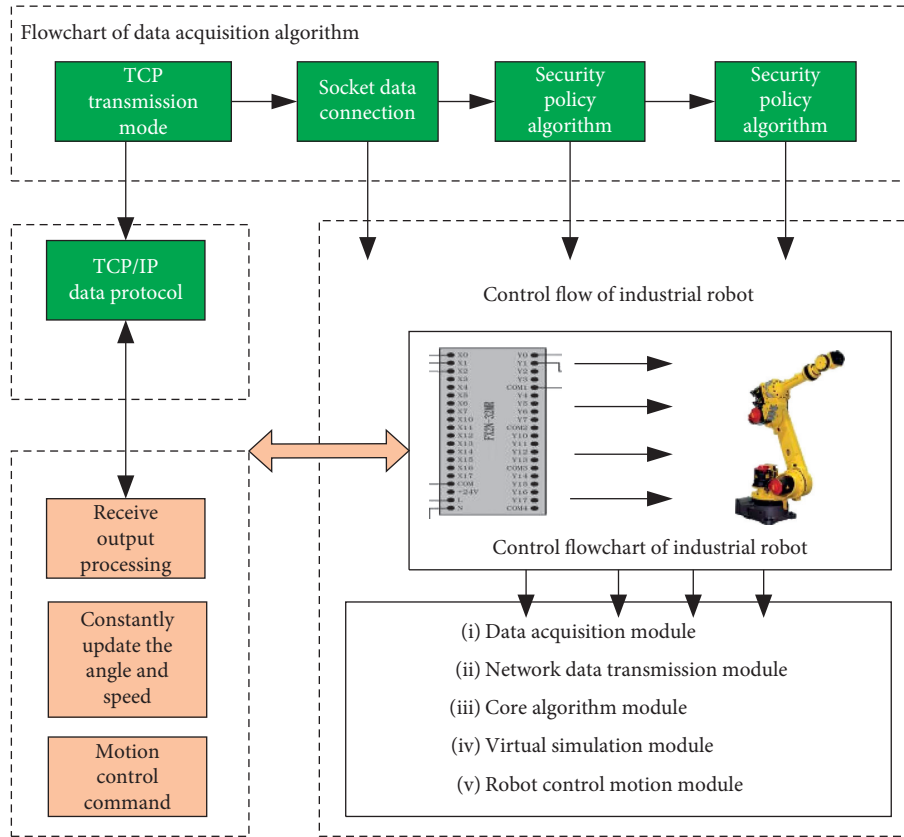
can be seen from the figure that the corresponding system includes the algorithm part and the hardware sensor part. The data acquisition module corresponding to the algorithm part is mainly responsible for collecting the real-time motion coordinate data of human and industrial robot. The corresponding network transmission is mainly responsible for the data transmission between each module. The corresponding planning algorithm is mainly responsible for collision detection between industrial robot and human, and it also needs to further calculate the next motion angle and speed of industrial robot in this case. The corresponding real-time calculation data in the core control module will be verified by the virtual simulation and robot control module, and the real-time motion control of industrial robot will be completed.

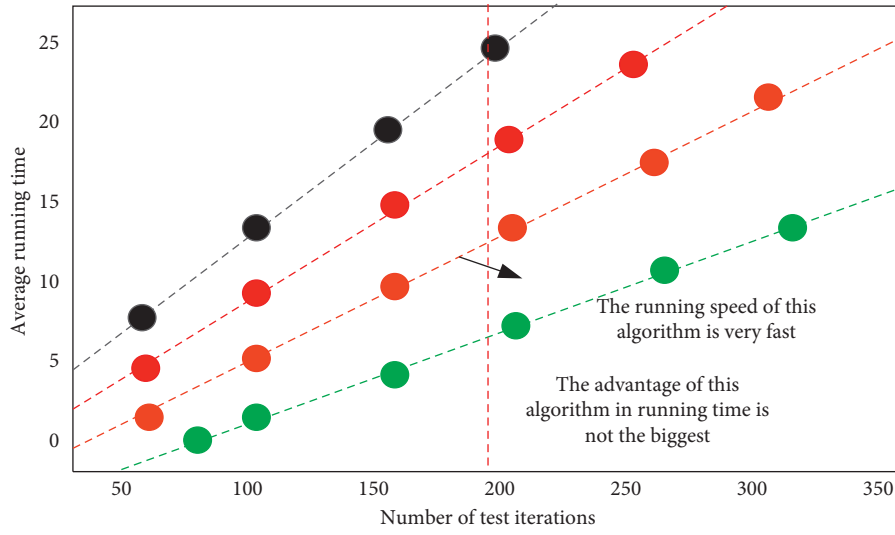
The corresponding hardware module design level mainly includes five hardware modules, which are data acquisition module, network data transmission module, core algorithm module, virtual simulation module, and robot control movement module. In the corresponding data acquisition module, it mainly collects the spatial coordinate information of industrial robot and human movement and sends the data collected to the corresponding analysis module of the algorithm. The corresponding data needs to be cached in the corresponding local species after the acquisition. In the actual transmission, the latest data is obtained from the local at a certain time and transmitted to the collision avoidance path planning program for processing. In the corresponding network data transmission module, it is mainly responsible for the transmission of data in each module of the system, so as to ensure the reliability of data. The corresponding transmission mode adopts TCP transmission mode and socket to connect different data, so as to realize the reliable

transmission of data. The corresponding algorithm module includes security strategy algorithm and path planning algorithm. It calculates the joint angle combination and corresponding motion speed of the next motion of industrial robot continuously and carries out real-time path planning processing based on the current corresponding data. At the corresponding speed control level of industrial robots, computer control line is mainly used to connect with the corresponding speed controller of industrial robot. The corresponding working flowchart of industrial robot control module is shown in Figure 7.

Based on the design of the above system, this paper carries out experimental simulation based on an industrial robot in the corresponding scene. The main simulation experiments of this paper include security strategy simulation and path planning simulation.

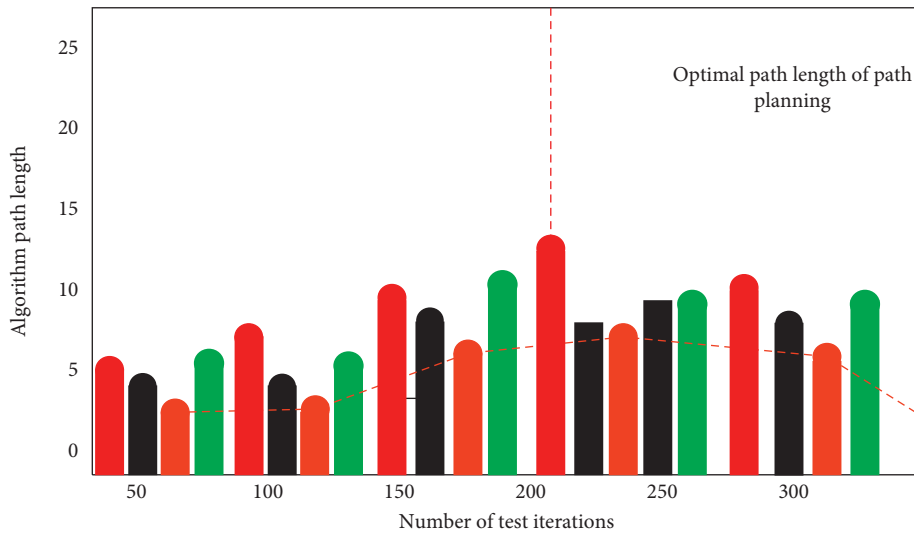
4.1. Security Policy Simulation Experiment. In order to further verify the obstacle avoidance ability of the industrial robot in the presence of environmental interference, on the premise that other control environment variables remain unchanged, observe the curve change of the end point of the industrial robot in the z -axis of the spatial coordinate in the interference and noninterference environment, and the corresponding change trajectory is shown in Figure 8. It can be seen from Figure 8 that, in the noninterference environment, the industrial robot has been moving along the direction corresponding to the z -axis and finally reaches the end point. In the corresponding interference environment, the industrial robot starts to avoid obstacles when it moves to the 34th step, and its corresponding joint conversion angle is also smaller and smaller. After successfully





--- IGA
 --- The algorithm in this paper
 --- IARF
 --- BI-RRT

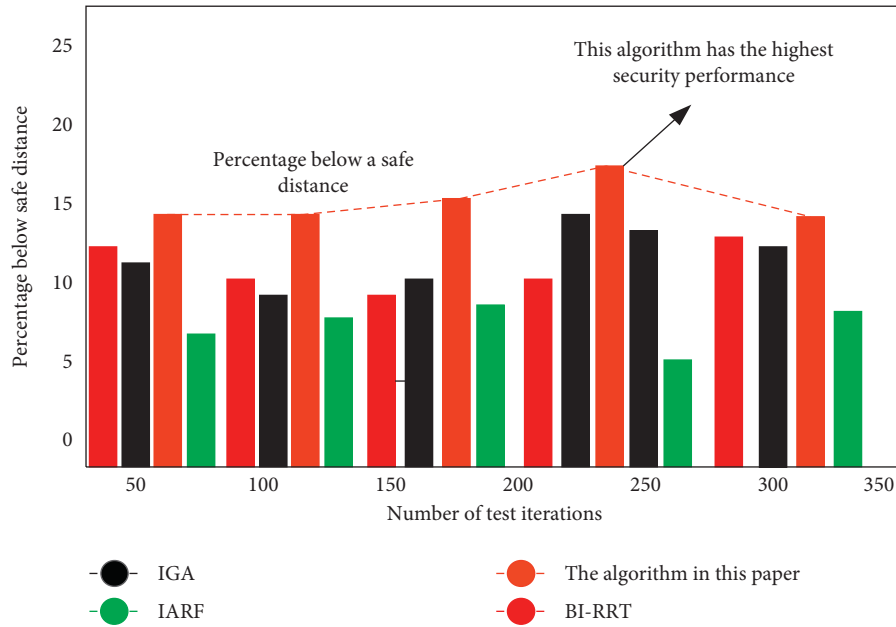
(a)



● IGA
 ● The algorithm in this paper
 ● IARF
 ● BI-RRT

(b)

FIGURE 9: Continued.



(c)

FIGURE 9: (a) Comparison curve of average running time of algorithm. (b) Comparison curve of path length of algorithm planning. (c) Algorithm security comparison curve.

be seen that the corresponding planning path of the algorithm proposed in this paper is the shortest and the corresponding algorithm is stable. Although the traditional algorithm has some advantages in time, the corresponding planning path is longer. Figure 9(c) shows the contrast curve of the algorithm and the traditional algorithm in the algorithm security. From the figure, it can be seen that the proposed algorithm can ensure that the distance between the links and obstacles of industrial robots is greater than the safety distance compared with the traditional algorithm, while the distance corresponding to other traditional algorithms will be less than the safety distance. This will cause the collision between industrial robot and obstacles or human beings, which will affect the safety strategy of the whole system.

Based on the above experiments and analysis, it can be concluded that the algorithm in this paper has obvious advantages compared with the traditional algorithm in the aspect of human-computer cooperation and path planning of industrial robot, and its corresponding obstacle avoidance effect is significant, which has high practical promotion value.

5. Summary

This paper mainly analyzes the development trend of industrial robot and the research hotspots of human-computer cooperation and path planning and gives the development difficulties of human-computer cooperation and path planning of industrial robot. In view of the above difficulties, this paper studies and analyzes the human-machine cooperation and path optimization of industrial robot in a complex road environment. Firstly, based on the kinematics

of industrial robot, the theoretical modeling and obstacle avoidance analysis of industrial robot are carried out; thus, an efficient and concise human-machine collision detection model of industrial robot is proposed. In view of the complex road conditions faced by industrial robots, this paper studies and analyzes the obstacle avoidance strategy of human-computer cooperation and real-time path optimization algorithm of industrial robots. Based on the virtual target point algorithm, it further improves the problem that the traditional path planning algorithm can not fully cover the target and puts forward the corresponding improved path planning algorithm of industrial robots. In the experimental part, based on the existing industrial robot system, this paper designs the human-computer cooperation and path planning system. The experimental results show that the algorithm proposed in this paper can achieve the overall obstacle avoidance of industrial robot more efficiently than the traditional algorithm, and the planned path is more safe and efficient, which verifies the feasibility and superiority of the algorithm. In the following research, this paper will focus on the application of the proposed algorithm in large-scale industrial scenes and study the corresponding algorithm consumption and convergence problems.

Data Availability

Relevant data can be obtained from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] F. Mars, M. Deroo, and J.-M. Hoc, "Analysis of human-machine cooperation when driving with different degrees of haptic shared control," *IEEE Transactions on Haptics*, vol. 7, no. 3, pp. 324–333, 2014.
- [2] Z. Yahouni, N. Mebarki, F. Belkadi, A. Shahzad, and A. Bernard, "Human-machine cooperation in planning and scheduling: a case study on an unstable environment," *European Journal of Industrial Engineering*, vol. 12, no. 6, pp. 757–764, 2018.
- [3] L. Athanasopoulou, A. Papacharalamopoulos, P. Stavropoulos, and D. Mourtzis, "Design and manufacturing of a smart mobility platform's context awareness and path planning module: a PSS approach," *Procedia Manufacturing*, vol. 51, no. 1–2, pp. 61–66, 2020.
- [4] X. L. Jia, X. Q. Yang, and Y. S. Deng, "Path planning of multi-robot fish based on cooperative game model," *Applied Mechanics and Materials*, vol. 341–342, no. 3, pp. 940–944, 2013.
- [5] E. Masehian, M. Jannati, and T. Hekmatfar, "Cooperative mapping of unknown environments by multiple heterogeneous mobile robots with limited sensing," *Robotics and Autonomous Systems*, vol. 87, no. 6, pp. 188–218, 2017.
- [6] J. Yang, D. Wang, B. Fan et al., "Online absolute pose compensation and steering control of industrial robot based on six degrees of freedom laser measurement," *Optical Engineering*, vol. 56, no. 3, pp. 034111.1–034111.9, 2017.
- [7] Y. Chao, X. Chen, and N. Xiao, "Deep learning-based grasp-detection method for a five-fingered industrial robot hand," *IET Computer Vision*, vol. 13, no. 1, pp. 61–70, 2018.
- [8] R. Martínez and J. Carlos, "A function block based approach using assembly features for controlling virtual and real industrial robots," *The Journal of Symbolic Logic*, vol. 30, no. 2, pp. 248–249, 2015.
- [9] S. S. Pchelkin, A. S. Shiriaev, A. Robertsson et al., "On orbital stabilization for industrial manipulators: case study in evaluating performances of modified PD+ and inverse dynamics controllers," *IEEE Transactions on Control Systems Technology*, vol. 11, no. 4, pp. 1–17, 2017.
- [10] A. Ferrara and G. P. Incremona, "Design of an integral suboptimal second-order sliding mode controller for the robust motion control of robot manipulators," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 6, pp. 2316–2325, 2015.
- [11] M. P. Polverini, S. Formentin, L. Merzagora et al., "Mixed data-driven and model-based robot implicit force control: a hierarchical approach," *IEEE Transactions on Control Systems Technology*, vol. 8, no. 99, pp. 1–14, 2019.
- [12] V. Nguyen, J. Johnson, and S. Melkote, "Active vibration suppression in robotic milling using optimal control," *International Journal of Machine Tools and Manufacture*, vol. 152, no. 5, Article ID 103541, 2020.
- [13] J. Baek, W. Kwon, B. Kim, and S. Han, "A widely adaptive time-delayed control and its application to robot manipulators," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 7, pp. 5332–5342, 2019.
- [14] B. Xiao, H. K. Lam, G. Song, and H. Li, "Output-feedback tracking control for interval type-2 polynomial fuzzy-model-based control systems," *Neurocomputing*, vol. 242, no. 12, pp. 83–95, 2017.
- [15] H. Hassan, C. Dominguez, J.-M. Martinez, A. Perles, J.-V. Capella, and J. Albaladejo, "A multidisciplinary PBL robot control project in automation and electronic engineering," *IEEE Transactions on Education*, vol. 58, no. 3, pp. 167–172, 2015.
- [16] T. Tsuji, J. Ohkuma, and S. Sakaino, "Dynamic object manipulation considering contact condition of robot with tool," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 3, pp. 1972–1980, 2016.
- [17] J. Qin, S. Wang, Y. Kang, and Q. Li, "Circular formation algorithms for multiple nonholonomic mobile robots: an optimization-based approach," *IEEE Transactions on Industrial Electronics*, vol. 15, no. 5, pp. 22–32, 2018.
- [18] J. Edwards, "Signal processing helps put robot users in control special reports," *IEEE Signal Processing Magazine*, vol. 33, no. 3, pp. 8–11, 2016.
- [19] J. Yu, W. Chen, and G. Xie, "Coordination of multiple robotic fish with applications to underwater robot competition," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 2, pp. 1–12, 2015.
- [20] P. Lu, W. Yu, G. Chen, and X. Yu, "Leaderless consensus of ring-networked mobile robots via distributed saturated control," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 12, pp. 10723–10731, 2020.
- [21] S. Rahman and R. Ikeura, "Weight-prediction-based predictive optimal position and force controls of a power assist robotic system for object manipulation," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 9, pp. 11–19, 2016.
- [22] X. Li and C. C. Cheah, "A simple trapping and manipulation method of biological cell using robot-assisted optical tweezers: singular perturbation approach," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 2, pp. 1656–1663, 2017.
- [23] C. Faria, C. Vale, T. Machado et al., "Experiential learning of robotics fundamentals based on a case study of robot-assisted stereotactic neurosurgery," *IEEE Transactions on Education*, vol. 59, no. 2, pp. 45–51, 2016.
- [24] A. H. Memar and E. T. Esfahani, "A robot gripper with variable stiffness actuation for enhancing collision safety," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 8, pp. 6607–6616, 2020.
- [25] H.-K. Shin and B. K. Kim, "Energy-efficient gait planning and control for biped robots utilizing vertical body motion and allowable ZMP region," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2277–2286, 2015.

Research Article

Adaptive Adjustment Object Detection Algorithm under Multiple Mechanisms Based on GAN

Zemin Qiu , Feng Wang , and Zhihong Pan

School of Information Science, Xinhua College of Sun Yat-Sen University, Guangzhou, China

Correspondence should be addressed to Zemin Qiu; qiuzemin@xhsysu.edu.cn

Received 16 April 2021; Revised 13 May 2021; Accepted 31 May 2021; Published 19 June 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Zemin Qiu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Target tracking is prone to problems such as target loss and identity jump when the target is occluded and the attitude is changed. In order to solve this phenomenon, this paper proposes the adaptive adjustment object detection algorithm under multiple mechanisms based on GAN. This algorithm introduces a gradient penalty mechanism to the discriminator and uses the relative discriminator structure to reconstruct the discriminator, so as to improve the discriminatory ability of the discriminator. Then, through the feedback mechanism, the obtained data is fed back to the generator network in time, and the genetic mechanism is used to speed up the positioning of the key areas of the image. Experimental results show that compared with other existing algorithms, this algorithm can effectively locate and distinguish under different environments. And, the target still maintains a high resolution. When the target is occluded, it can effectively avoid the phenomenon of target loss and identity jump.

1. Introduction

Target detection is one of the important research directions in the field of computer vision. It integrates technologies such as positioning, recognition, and classification and is widely used in fields such as intelligent monitoring and human-computer interaction. In recent years, the target detection algorithm has made great breakthroughs. The representative target detection algorithms can be divided into two main categories. One is the R-CNN algorithm (R-CNN, Fast R-CNN, Faster R-CNN, etc.) [1, 2]. The algorithm needs to generate the target first. Candidate frame (target location), and then, classify and regress the candidate frame. The other is one-stage algorithms such as Yolo and SSD [3–5], which use the convolutional neural network (CNN) to extract image features and establish detection models to classify and locate targets. After comparison, the second type of algorithm is faster, but the accuracy is lower. In practical applications, the target in the image often has problems such as being occluded and deformed, making the extracted target feature insignificant or lacking part of the feature information, leading to classification errors and tracking failures,

and then prone to the problem of identity jump. Therefore, we are comparing the advantages and disadvantages of existing algorithms; this paper adopts the generative confrontation network in the second type of algorithm. This model is a unique network structure that captures potential data distribution. The true and false confrontation error design based on game theory construction makes it possible to deal with different types of tasks [6, 7].

Based on the idea of generative confrontation, this paper proposes a multimechanism adaptive target detection algorithm. First, through the organic combination of the punishment mechanism, feedback mechanism, and scoring mechanism, it effectively solves the problems of difficulty in training and model collapse of the generative confrontation network. Then, use the relative discriminator structure to construct the discriminator to improve the probability of the discriminator in the process of fighting competition. The result is sent to the generator network through the feedback mechanism, and finally, the individual with the optimal value is obtained through the scoring mechanism, so as to quickly and effectively locate the key area of the image. After experiments, the experimental results are basically in line

with expectations. The target detection algorithm can effectively locate and distinguish different objects and still maintain a high resolution for the occluded target, while maintaining a low number of missing alarms and false alarms.

2. Generative Adversarial Network (GAN)

The generative adversarial network (GAN) was proposed by Goodfellow. It uses two convolutional neural networks using game training to generate samples similar to the original picture [8–10].

By observing the structure of the generative confrontation network in Figure 1, GAN is mainly composed of two parts: generation network (G) and discriminant network (D). And, it uses a competing mechanism [11–13]; the purpose is to enable the generator to generate data similar to the real data distribution, which made the GAN achieve the effect of being fake. Aiming at the problem of data prediction, the generator is trained to learn the data distribution and predict the image distribution data. Generated data distribution and real data distribution are shown in Figure 2.

Because the difference between the generated data distribution P_G and the real data distribution P_{data} cannot be calculated, therefore, the discriminator uses data sampling to calculate the divergence of P_G and P_{data} . And, use these to discriminate the difference between the two data distributions, and the objective function of the discriminator is

$$V(G, D) = E_{x \sim P_{\text{data}}} [\log D(x)] + E_{x \sim P_G} [\log(1 - D(x))]. \quad (1)$$

Among them, $E(*)$ is mathematical expectation and $D(*)$ is the output result of the discriminator. The objective function is to maximize the discriminator, equivalent to the divergence generated data and real data. The optimal generator achieves the purpose of prediction by minimizing the divergence of the distribution between the real image data and the generated data. The optimal generator can be expressed as

$$G^* = \arg \min_G \max_D V(G, D). \quad (2)$$

GAN can generate samples that are similar to the real data distribution, but it has problems such as difficulty in training, difficulty in network convergence, and mode collapse.

3. The Discriminator Network Introduces a Gradient Penalty Mechanism

In order to make the training of the model more stable and solve the problem of model collapse, this paper adds a gradient penalty to the discriminator objective function. After adding the gradient penalty, the objective function is

$$V(G, D) = E_{x \sim P_{\text{data}}} [D(x)] - E_{x \sim P_G} [D(x)] - \lambda_{gp} E_{x \sim P_{\text{penalty}}} \left[\left(\|\nabla_x D(x)\|_2 - 1 \right)^2 \right]. \quad (3)$$

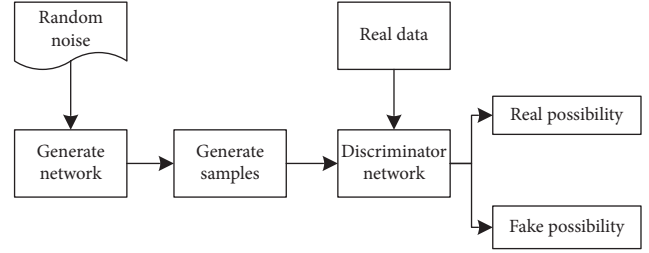


FIGURE 1: Generative adversarial network structure.

Among them, λ_{gp} is the weight parameter. P_{penalty} is the probability distribution of the sample for calculating the gradient penalty. Sampling a point in P_G and P_{data} , respectively, the method is random sampling. $\|\cdot\|_2$ is a two-norm number.

In the proposed data prediction algorithm, the loss function can reduce the difference in probability distribution between the predicted data and the real data so that the predicted data generated by the generator can more accurately fit the real data distribution. The loss function is defined as

$$L(G, D) = E_{x \sim P_{\text{data}}} [D(x)] - E_{x \sim P_G} [D(x)] - \lambda_{gp} E_{x \sim P_{\text{penalty}}} \left[\left(\|\nabla_x D(x)\|_2 - 1 \right)^2 \right]. \quad (4)$$

Adding a traditional loss function to the loss function of GAN can improve the accuracy of the data generated by the generator. The role of the generator is not only to generate data that can deceive the discriminator, but the generated data should be close enough to the real data. Since data prediction is a regression problem, adding the output result data of the discriminator to the loss function can punish the extreme error between the predicted value and the true value. Therefore, the loss function after introducing the gradient penalty is

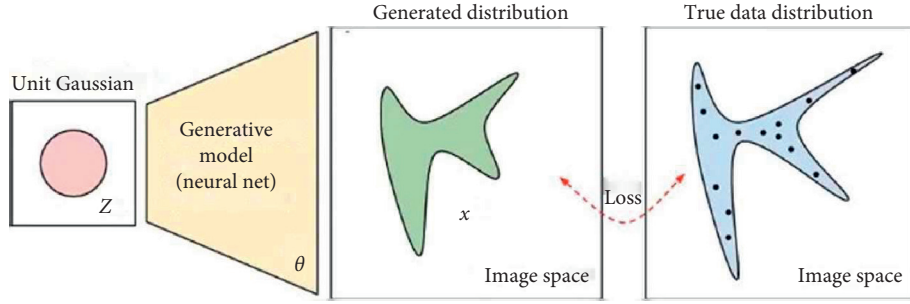
$$L_{\text{MSE}}(G, D) = L(G, D) + \lambda_{\text{MSE}} L_{\text{MSE}}(G), \quad (5)$$

where λ_{MSE} is the weight parameter, and the loss function of $L_{\text{MSE}}(G)$

$$L_{\text{MSE}}(G) = E_{x \sim P_{\text{data}}} [\|x - G(x)\|_2]. \quad (6)$$

Through research, it can be found that $L_{\text{MSE}}(G)$ is the loss function that is only extracted for the generator. Just considering the generator data to improve the accuracy of the algorithm is not enough. In GAN, the data generated by the discriminator corresponds to the generator has a guiding role.

Therefore, this paper adopts a feedback mechanism for the data output by the discriminator network based on the introduction of a gradient penalty mechanism in the generator. This improves the accuracy of the loss value $L_{\text{MSE}}(G)$. The discriminator in this article is constructed with a relative discriminator structure to estimate the probability that the real image is more real than the fake image and then promote the discriminator to obtain better probability results in the process of fighting competition, thereby increasing the

FIGURE 2: Generated data distribution P_G and real data distribution P_{data} .

loss of the generator value $L_{\text{MSE}}(G)$ to strengthen the network parameter optimization of the generator.

By collecting the real data and false data of the discriminator and estimating the corresponding sample expectations, the following function is obtained:

$$\begin{aligned} D_{\text{RI}}(I_{\text{real}}, I_{\text{fake}}) &= \sigma\{D(I_{\text{real}}) - E_{I_{\text{fake}}}(D(I_{\text{fake}}))\}, \\ D_{\text{RI}}(I_{\text{fake}}, I_{\text{real}}) &= \sigma\{D(I_{\text{fake}}) - E_{I_{\text{real}}}(D(I_{\text{real}}))\}. \end{aligned} \quad (7)$$

Among them, $E_{I_{\text{real}}}$ and $E_{I_{\text{fake}}}$ are the sample expectations of real data and fake data, respectively. This paper uses the minimum loss function to calculate.

For the global discriminator, the ideal values of $E_{I_{\text{fake}}}(D(I_{\text{fake}}))$ and $E_{I_{\text{real}}}(D(I_{\text{real}}))$ should be close to 0 and 1. The expected discrimination result is $D(I_{\text{real}}) = 1$ and $D(I_{\text{fake}}) = 0$. Therefore, the regression targets of $D_{\text{RI}}(I_{\text{real}}, I_{\text{fake}})$ and $D_{\text{RI}}(I_{\text{fake}}, I_{\text{real}})$ are 1 and 0.

For the generator, its role is to confuse the authenticity judgment of the discriminator as much as possible, which is opposite to the return goal of the discriminator. The regression targets of $D_{\text{RI}}(I_{\text{real}}, I_{\text{fake}})$ and $D_{\text{RI}}(I_{\text{fake}}, I_{\text{real}})$ are 0 and 1. Therefore, the adversarial function to obtain the global discriminator is defined as follows:

$$\begin{aligned} L_{\text{MSE}}(G) &= \frac{1}{2}E_{I_{\text{fake}}}\{D_{\text{RI}}(I_{\text{fake}}, I_{\text{real}})^2\} \\ &+ \frac{1}{2}E_{I_{\text{real}}}\{[D_{\text{RI}}(I_{\text{real}}, I_{\text{fake}}) - 1]^2\}. \end{aligned} \quad (8)$$

In this paper, by setting the feature-preserving loss function, the confrontation function $L_{\text{MSE}}(G)$ of the global discriminator is input into the above loss function for feature error calculation; finally, the loss function is defined as follows:

$$L(G, D) = E_{x \sim P_{\text{data}}}[D(x)] - E_{x \sim P_G}[D(x)] - \lambda_{gp} E_{x \sim P_{\text{penalty}}}\left[\left(\|\nabla_x D(x)\|_2 - 1\right)^2\right], \quad (9)$$

$$L_{\text{MSE}}(G, D) = L(G, D) + \lambda_{\text{MSE}}\left(\frac{1}{2}E_{I_{\text{fake}}}\{D_{\text{RI}}(I_{\text{fake}}, I_{\text{real}})^2\} + \frac{1}{2}E_{I_{\text{real}}}\{[D_{\text{RI}}(I_{\text{real}}, I_{\text{fake}}) - 1]^2\}\right).$$

By observing Figures 3–5, it can be seen that, after the reconstruction of the generated network in this paper, the image is basically effectively repaired, and it has a good real consistency. Although the repaired part is somewhat unclear and may have subtle changes, but it basically does not affect the detection effect.

4. Generating Network Scoring Mechanism

However, for the targets in the video, they are basically in a state of movement and overlap. We need to quickly and effectively locate the key areas of the image [14, 15]. Therefore, based on the above formula, this paper proposes a

scoring mechanism for the generative network in the generative confrontation network. Each iteration of the generator network is composed of three stages: mutation, evaluation, and selection so that the generator can quickly and accurately locate the key areas in the image.

4.1. Variation. Mainly for the input random noise signal z , the gradient penalty loss function of this article is used in the new round of update process. The penalty loss function $L(G, D)$ for the generating network gradient is calculated only for noise data, and finally, $L_{\text{MSE}}(G, D)$ is obtained and the network data is updated:

$$L_z(G, D) = E_{z \sim P_z}[D(z)] - E_{z \sim P_z}[D(z)] - \lambda_{gp} E_{z \sim P_z}\left[\left(\|\nabla_z D(z)\|_2 - 1\right)^2\right], \quad (10)$$

$$L_{\text{MSE}}(G, D) = L_z(G, D) + \lambda_{\text{MSE}}\left(\frac{1}{2}E_{I_{\text{fake}}}\{D_{\text{RI}}(I_{\text{fake}}, I_{\text{real}})^2\} + \frac{1}{2}E_{I_{\text{real}}}\{[D_{\text{RI}}(I_{\text{real}}, I_{\text{fake}}) - 1]^2\}\right).$$



FIGURE 3: Original image (occluded).



FIGURE 4: Generated image.

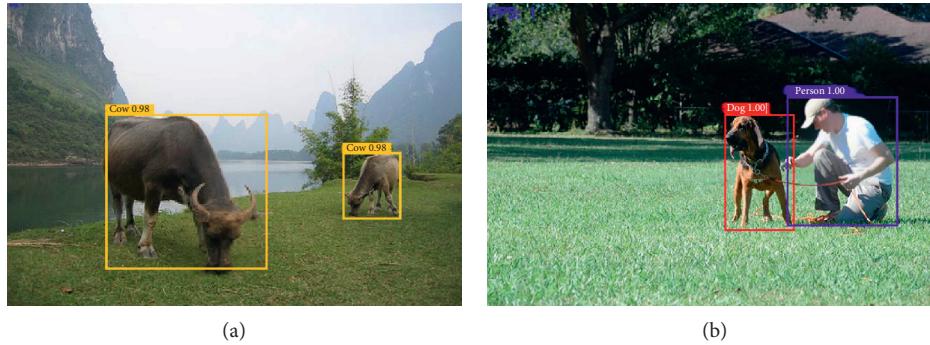


FIGURE 5: Object detection.

Among them, λ_{gp} is the weight parameter and P_z is the distribution of noise data.

4.2. Evaluation. This stage is the data evaluation stage. The network uses the above loss function and selects the best sample under the feedback of the discriminator, so as to learn the correct training sample distribution and improve the stability of model training. The evaluation mechanism is as follows.

This stage is quantified as an adaptability score, which mainly analyzes the quality of the samples generated by the network and the diversity of the generated samples. It evaluates the quality of the generated samples and the

diversity of the generated samples; the adaptability score can be expressed as

$$\mathbf{F} = \mathbf{F}_q + \gamma \mathbf{F}_d. \quad (11)$$

Among them, \mathbf{F}_q is the quality score of the generated sample and \mathbf{F}_d is the diversity score of the generated sample. It is used to measure whether the generator can fully disperse the generated samples, which can further avoid mode collapse. γ is a super parameter that balances the proportion of the \mathbf{F}_q and \mathbf{F}_d .

Calculate the highest fitness score. And, compare it with the highest score on the network to get the global best score.

4.3. Selection. Select the best individual according to the evaluation mechanism. Others with lower scores will

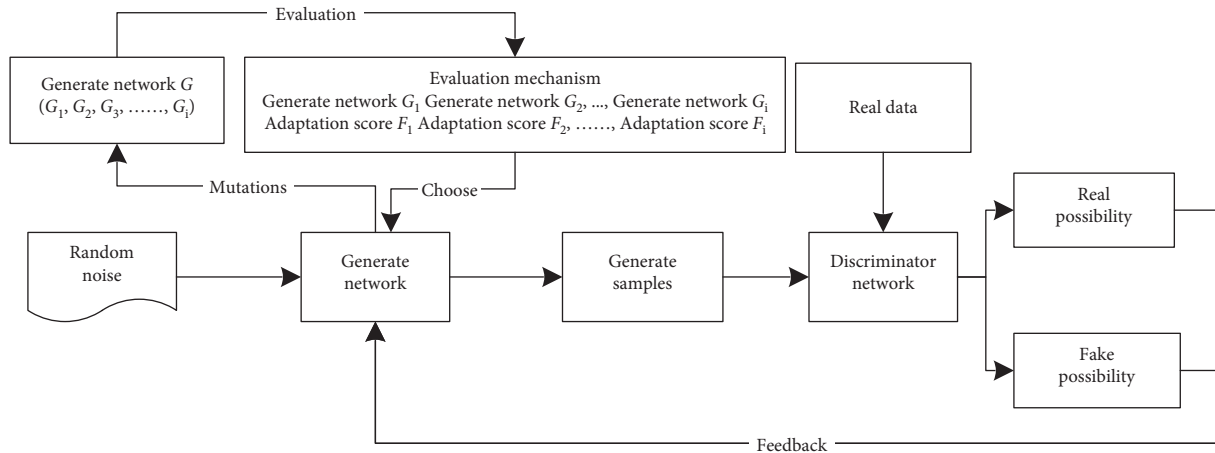


FIGURE 6: The structure of the article generation adversarial network.



FIGURE 7: Effect picture of object detection.

gradually decay so that the generated network has a certain directionality, and the best value sample is selected.

The structure of the article generation adversarial network is shown in Figure 6.

5. Simulation Experiment

This paper uses the ImageNet dataset and the MOT training dataset to test the algorithm. All experiments in this paper are carried out in the same experimental environment, and

TABLE 1: Evaluation criteria in MOTChallenge.

Measure	Better	Perfect	Description
MOTA	Higher	100%	Tracking accuracy: it is negatively correlated with the number of FN, FP, and IDs
MOTP	Higher	100%	Tracking accuracy, matching of GT, and detected bbox overlap
IDF1	Higher	100%	F1 introducing track ID
FP	Lower	0	The total number of FPs; false positives are false positives
FN	Lower	0	The total number of FNs; false negatives are also missed
IDs	Lower	0	Total number of ID changes
Hz	Higher	Inf	Processing speed, excluding the time consumption of the detector

the data obtained by the article algorithm are compared with traditional algorithms, YOLO-based target detection, Deep Sort, and literature algorithms [16] to verify the effectiveness of the algorithm.

Considering that there is a big difference between artificial occlusion and occlusion in the real world, this chapter directly simulates the actual dataset, and the result of target detection is shown in the figure below.

By observing the target detection effect map in Figure 7, we can see that the article algorithm can effectively locate and distinguish different objects and still maintain a high resolution for the occluded target, even if the characters in Figures 7(d) and 7(i) are occluded. More than half of the area can still accurately identify and mark the target.

5.1. Accuracy Test. This paper uses the evaluation criteria in MOTChallenge to evaluate the performance of the proposed algorithm and compares the measured data with traditional target detection algorithms and literature algorithms. Among them, the MOTChallenge dataset has a comprehensive data type, which includes different viewing angles, camera movement methods, and pictures taken in different climates, which can effectively evaluate algorithms. Standards are shown in Table 1.

This article uses evaluation criteria: MOTA, MOTP, IDs, FP, and FN. By observing the identity jump problem of the algorithm in the target detection process. Through the test, we can obtain the data as shown in Table 2.

By observing the target tracking effect data comparison table in Table 2, we can see that the MOTA, IDs, and FP in this paper are better than other algorithms. In terms of MOTP data, the article algorithm is only 0.4% higher than the literature algorithm. It can be seen that the article algorithm can preserve better detection accuracy, while performing well in the number of false alarms (FP) and number of missed alarms (FN). It can effectively prove that the article algorithm can effectively reduce the problem of identity jumps.

By observing Figure 8, it can be seen that, under the same accuracy, the number of identity jumps of the algorithm in this paper is smaller than that of other algorithms. And, in terms of the highest accuracy, the article algorithm is about 2% higher than the literature algorithm. Observing Figure 9, one can clearly see that the

TABLE 2: Comparison table of target tracking effect data.

Algorithm	MOTA	MOTP	IDs	FP	FN
Traditional algorithm	35.9	62.5	923	14332	73122
YOLO object detection algorithm	70.7	70.9	936	8742	53386
Deep Sort	63.2	81.2	743	11261	62311
Document algorithm	71.2	82.6	257	4410	31001
Article algorithm	72.1	83.0	122	2738	34219

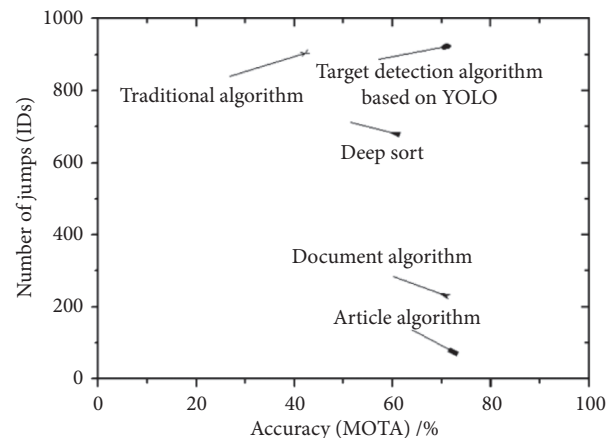


FIGURE 8: Comparison of accuracy and number of identity jumps.

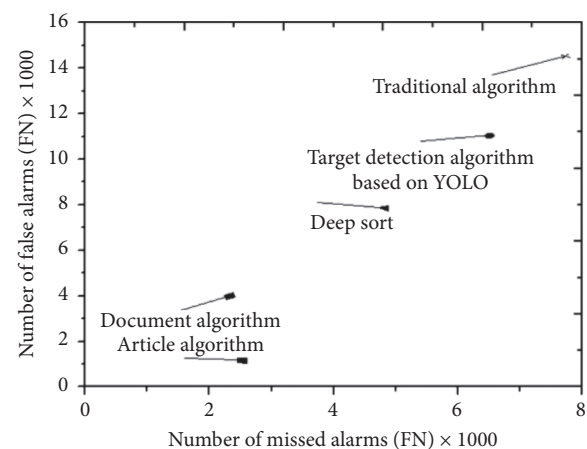


FIGURE 9: Comparison of the number of missed alarms and the number of false alarms.

algorithm in this paper is in the lower left corner of the image, and the number of missed alarms and false alarms are the lowest.

6. Conclusion

There are problems such as target loss and identity jump in current target detection. In response to these phenomena, we propose a GAN-based multimechanism adaptive target detection algorithm. The article optimizes the construction model of GAN by introducing gradient penalty and feedback and scoring mechanism into the original GAN, which achieves an effective improvement in detection accuracy. Experimental results show that compared with other existing algorithms, the article algorithm has a more stable training model, which can effectively reduce the problem of target identity jump, while maintaining high detection accuracy. Although the method is effective, there are still areas that can be improved, such as the use of a more appropriate size convolution kernel to extract features, which will continue to be studied in future works.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

The authors have equally contributed to the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This research was supported by the Scientific Research Platforms and Project of Colleges and Universities in Guangdong Province under Grant no. 2018KTSCX316, Science and Technology Project of Guangzhou under Grant no. 202002030273, and Discipline Project, Xinhua College of Sun Yat-sen University with no. 2020XZD02.

References

- [1] R. Yun, Z. Changren, and X. Shunping, "Object detection based on fast/faster RCNN employing fully convolutional architectures," *Mathematical Problems in Engineering*, vol. 2018, Article ID 3598316, 7 pages, 2018.
- [2] L. Zhang, Y. Zhang, and Z. Zhang, "Real-time water surface object detection based on improved faster R-CNN," *Sensors*, vol. 19, no. 16, p. 3523, 2019.
- [3] X. Zhang, X. Dong, Q. Wei et al., "Real-time object detection algorithm based on improved YOLOv3," *Journal of Electronic Imaging*, vol. 28, no. 5, p. 1, 2019.
- [4] L. Zhao and S. Li, "Object detection algorithm based on improved YOLOv3," *Electronics*, vol. 9, no. 3, pp. 1–11, 2020.
- [5] C. Sun, Y. Ai, S. Wang et al., "Mask-guided SSD for small-object detection," *Applied Intelligence*, vol. 51, no. 20, pp. 1–12, 2020.
- [6] Z. Wu, C. He, L. Yang et al., "Attentive evolutionary generative adversarial network," *Applied Intelligence*, vol. 51, no. 6, pp. 1–15, 2020.
- [7] Q. Cui, H. Sun, Y. Kong, X. Zhang, and Y. Li, "Efficient human motion prediction using temporal convolutional generative adversarial network," *Information Sciences*, vol. 545, pp. 427–447, 2021.
- [8] B. Jiang, Z. Zhou, X. Wang et al., "cmSalGAN: RGB-D salient object detection with cross-view generative adversarial networks," *IEEE Transactions on Multimedia*, vol. 1, pp. 1–11, 2020.
- [9] J. Li, C. Qu, S. Peng, and Y. Jiang, "Ship detection in SAR images based on generative adversarial network and online hard examples mining," *Journal of Electronics & Information Technology*, vol. 41, no. 1, pp. 143–149, 2019.
- [10] G. Chen, L. Liu, J. Guo, Z. Pan, and W. Hu, "Semi-supervised airplane detection in remote sensing images using generative adversarial networks," *Journal of University of Chinese Academy of Sciences*, vol. 37, no. 4, pp. 539–546, 2020.
- [11] C. Wang and J. Tian, "Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks," *CAAI Transactions on Intelligent Systems*, vol. 15, no. 2, pp. 296–301, 2020.
- [12] X. Zhai, Z. Cheng, Y. Wei et al., "Compressive sensing ghost imaging object detection using generative adversarial networks," *Optical Engineering*, vol. 58, no. 1, p. 1, 2019.
- [13] Z. Liu, W. Zhang, and P. Zhao, "A cross-modal adaptive gated fusion generative adversarial network for RGB-D salient object detection," *Neurocomputing*, vol. 387, pp. 210–220, 2020.
- [14] L. Courtrai and M. T. Pham, "Small object detection in remote sensing images based on super-resolution with auxiliary generative adversarial networks," *Remote Sensing*, vol. 12, no. 19, p. 3152, 2020.
- [15] A. N. Shan, S.-k. Lin, J.-z. Qiao, and C.-h. Li, "Object detection via learning occluded features based on generative adversarial networks," *Control and Decision*, vol. 1, pp. 1–7, 2020.
- [16] J. Li, X. Duan, M. Xu, and G. Xue, "Salient object detection based on conditional generative adversarial networks for video," *Transducer and Microsystem Technologies*, vol. 38, no. 11, pp. 129–132, 2019.
- [17] Y. Wei, C.-q. Xu, Z.-f. Diao, and B.-q. Li, "A multi-target pedestrian tracking algorithm based on generated adversarial network," *Journal of Northeastern University (Natural Science)*, vol. 41, no. 12, pp. 1673–1679, 2020.

Research Article

Applied Research of Knowledge in the Field of Artificial Intelligence in the Intelligent Retrieval of Teaching Resources

XuJing Bai  and JiaJun Li

School of Management, Northwestern Polytechnical University, Xi'an 710072, China

Correspondence should be addressed to XuJing Bai; bxjblue@mail.nwpu.edu.cn

Received 1 April 2021; Revised 20 May 2021; Accepted 27 May 2021; Published 14 June 2021

Academic Editor: Chenxi Huang

Copyright © 2021 XuJing Bai and JiaJun Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the development process of education informatization, digital teaching resources continue to grow, and how to manage and organize massive teaching resources has become a key issue for teaching staff. An efficient and accurate search system is an important part of the teaching resource service system. The use of intelligent search engines can search teaching resources comprehensively and efficiently, and the artificial intelligence search engine provides a reliable and convenient solution for the design and development of intelligent search systems. The article analyzes the design principles and technical standards of the intelligent search system, expounds the system's functional architecture and database design, and introduces the realization process and principles of the search system. By analyzing the characteristics of basic education resources and existing automatic abstracting methods, this paper proposes to integrate the calculated feature word weights in the field of basic education into the algorithm for calculating the weights of abstract sentences and simultaneously examine the sentence position, sentence length, and other texts. There is an automatic summarization algorithm for surface statistics. This article also introduces the search design ideas and implementation steps based on artificial intelligence, makes a scientific evaluation and summary of the actual situation of the automatic abstract system running in the basic education resource search engine, and looks forward to the next improvement work.

1. Introduction

At present, all kinds of schools in our country attach great importance to the construction of teaching resource databases and course databases. They have independently developed a large number of teaching resource databases such as courseware resource databases, course resource databases, and resource websites [1]. Teachers and students often use these resource databases to obtain learning resources and participate in the Internet. Teaching activities can effectively improve the information collection and processing capabilities of teachers and students. However, there are still the following problems in the process of using: a large number of repeated constructions of teaching resources and the established teaching courseware, electronic teaching plans, and course videos are insufficient in classification and resource organization standardization and perfection. The

results obtained by learners are often similar or even the same information when searching. It takes a lot of time to perform manual secondary retrieval and filtering; learners have individualized needs for the acquisition of the expression of teaching resources [2]. It is necessary to quickly find the target resource through positioning and retrieval. However, the resource organization structure of various teaching resource libraries and resources is different or heterogeneous, the resource manifestation is single, and the storage format is diverse [3]. The manifestation of the acquired teaching resources cannot meet the needs of learners and makes the acquisition of teaching resources more difficult.

This paper analyzes the relevant knowledge and technology of search engines, especially after analyzing some limitations of search engines, and puts forward an intelligent search engine model suitable for students in the environment of teaching

resources. The architecture of the intelligent search engine is given. In this model, the semantic network and intelligent agent technology in artificial intelligence technology are introduced to improve the intelligence of search engines. According to the characteristics of teaching resources, based on the classification of educational disciplines, the concept of a semantic network is constructed to realize the expansion of students' query requests, including synonymous expansion, related associations, and conceptual expansion. In order to achieve the purpose of expanding the scope of retrieval and improving the recall rate of search engines, by analyzing the feedback information of students, a personal interest model of students is established on the structure of the conceptual semantic network, and the degree of relevance of the concepts is expressed through the connection between the concepts. The model can intuitively reflect the range of students' personal interests. By establishing the range of students' interests, the retrieved documents will be reordered, so as to improve the ranking of the documents that students are interested in and improve the accuracy of search engines. This paper also tested and verified the intelligent agent system and gave the test results. The results show that the intelligent agent system can effectively improve the ranking of related documents and provide students with educational resources that interest them.

This article mainly describes the content of the paper from five parts. The specific arrangements of each chapter are summarized as follows. The first section is the introduction, which mainly introduces the specific topic background and research significance of the paper and analyzes the artificial intelligence search model and educational resource retrieval. The status quo summarizes the main research content and the structure of the paper. The second section discusses related work. The third section introduces the application of artificial intelligence technology in the intelligent retrieval of teaching resources. The fourth section analyzes the design of educational resource search of the artificial intelligence search algorithm, and the fifth section verifies the performance of artificial intelligence search through experimental simulation. The sixth section summarizes and looks forward to this article.

2. Related Work

With the continuous growth of information on the Internet, people have become a way of learning to find what they need through the Internet [4]. In particular, for students in education, their learning is based on resource learning. Therefore, obtaining useful subject knowledge is a key issue for online education students to improve their knowledge. Artificial intelligence search has gradually become a research hotspot for scholars at home and abroad [5]. The application of artificial intelligence search knowledge has received extensive attention and development in recent years. The research on artificial intelligence search knowledge in intelligent search of teaching resources is as follows.

Related scholars have conducted research on the on-site search subsystem and Chinese intelligent search engine in the online teaching platform. The guiding ideology of online teaching activities is a learner-centered educational

philosophy, which aims to improve their own learning ability through learners' independent learning and collaborative learning. The emergence and use of online teaching platforms provide a platform for learners and teachers to share teaching resources [6]. However, due to the massive number of teaching resources in the online teaching platform, the content and form of courses, after-school exercises, reference materials, teaching courseware, etc, they have the characteristics of diversity, which makes them difficult for teachers and learners to quickly search and apply resources in teaching activities [7]. How to maximize the sharing of teaching resources so that learners can quickly and accurately find the resources they need from their own resources is the main task of the development of search engines in the education field, and the development of a platform that can be applied to online teaching the resource search tool in is the core to solve this problem.

Related scholars analyzed the characteristics of online teaching platforms and the characteristics of online education, combined with the evaluation standards of online learners' learning ability, confirmed the theoretical value of search tools in online teaching platforms, and fully analyzed the working principles of search engines, and on the basis of key technologies, a search framework based on open source is designed. The overall architecture of the search system on the site, the introduction of information push technology and intelligent knowledge base technology into the framework, has improved the intelligence level of search engines, in addition, analyzed and compared the shortcomings of the word segmentation of existing search engines, proposed a word segmentation optimization plan, and optimized the search result ranking technology, making the search tool more suitable for learners to search and query massive resources in the teaching platform, and can quickly locate the location of the data required by the user [8]. The search results of the required resources are fed back to the users to further improve the search quality of the search engine. In the process of online teaching, learners mainly obtain the information resources they need through online media [9]. Whether learners can make full use of the teaching resources on the Internet is the main factor that affects their knowledge construction. However, for the current messy network in terms of teaching resources, learners will inevitably produce the phenomenon of "information trek". These factors will indirectly affect learners' interest and enthusiasm for learning. For this reason, using search engines to assist online teaching can improve resource utilization [10]. Search engines can not only improve the utilization of online teaching resources but also be one of the key technologies that need to be used in the development of online teaching.

2.1. Application of Artificial Intelligence Technology in Intelligent Retrieval of Teaching Resources

2.1.1. Construction of Artificial Intelligence Retrieval Ontology Model. Ontology is a clear and standardized description of conceptual models, while domain ontology is an abstract description of concepts and relationships in specific

domains. Its functions are similar to relational models, and it is an organizational framework for information resources in related domains [11]. Domain ontology is the starting point of the entire intelligent information retrieval system, and it also runs through the entire system structure, provides references and basis for other functional modules, and plays a pivotal role in the entire system. The construction of domain ontology is a complex system engineering, which requires the participation of many domain experts and a lot of time investment [12]. Various ontology construction methods have been proposed. In order to ensure that a suitable domain ontology can be established, this paper fully draws on the ideas and experience of software engineering and proposes a new prototype ontology construction method based on the spiral model. The ontology construction process is shown in Figure 1.

2.2. Application of Intelligent Classification Technology in the Classification of Teaching Resources. There are many kinds of teaching resources and a large number of them. Most of the data and information contained in them are in disorder. These data, information materials, and knowledge are presented in different formats and languages, and the source of each information resource is also different. In order to find accurate information, data, and knowledge as soon as possible, it is necessary to do a good job of processing information in advance [13]. The use of intelligent classification technology in archive classification management can effectively manage the messy knowledge, information, and data in the intelligence of teaching resources. According to certain standards, the huge and disorderly knowledge, information, and data can be classified in time to facilitate the location and acquisition of information data [14]. When categorizing, you can first divide the knowledge, information, and data into several major topics according to the theme's standards and then divide each major topic into several different small topics to construct a hierarchical topic structure. Teaching resource intelligence can be intelligently classified with the support of artificial intelligence technology, pattern recognition, natural language processing, and machine learning. In the technology of natural language processing, the text classification technology is based on a practice text sample library, and this sample library has been marked. In this sample library, a model that is related to text types and text attributes is found, and then, the found relational model is used to distinguish new text categories. The intelligence of text-based teaching resources can be intelligently classified with the support of text classification technology [15]. In addition, archivists can also intelligently distinguish and classify the teaching resources of multimedia categories by intelligent classification technology composed of voice recognition, image recognition, and video recognition. The system structure of intelligent search is shown in Figure 2.

2.3. Application of Intelligent Search Technology in Intelligent Retrieval of Teaching Resources. With the continuous development of artificial intelligence technology, the types and

quantity of teaching resource intelligence have grown rapidly, which has had a huge impact on traditional archival information resource retrieval [16]. The disadvantages of traditional retrieval methods are obvious, and retrieval efficiency has fallen far short of the requirements of archival work. Therefore, in order to improve the efficiency of archive retrieval, it is necessary to use intelligent retrieval technology to replace traditional archive retrieval methods. Intelligent retrieval makes full use of various artificial intelligence technologies such as pattern recognition and natural language processing [17]. The biggest difference between intelligent search and traditional search is as follows: intelligent search can search for the most similar results to the search content entered by the user and can also arrange these search results in the most similar order, which can greatly reduce the search time for users and improve retrieval efficiency. In addition to searching for text information, intelligent retrieval can also search for multimedia information such as sound, image, and video. In the current Internet age where multimedia teaching resources are more and more intelligent, intelligent retrieval technology has the advantages that traditional information retrieval technology cannot compare and has played a huge role in archive retrieval work [18]. The most fundamental purpose of file classification management is to maximize the use of information, data, and knowledge in the intelligence of teaching resources and to manage them efficiently. The intelligence, data, and knowledge of teaching resources have a wide variety and huge amount, and the sources of information, data, and knowledge are also different, which creates obstacles for users to search for the useful information [19]. Therefore, in order to meet the needs of users for archive search and want users to find the information they need in the shortest time, new search technologies must be introduced. Neural network technology is an artificial intelligence technology that imitates the behavioral characteristics of animal neural networks to perform distributed processing of information. In the file classification management work, the comprehensive introduction of neural network technology can greatly improve the searchability of search engines. As long as an intelligent search engine is developed, an accurate search of information, data, and knowledge in the intelligence of teaching resources can be realized, and the problem of information, data, and knowledge overload can be solved, so as to realize the effective management of archives classification [20]. The application of the intelligent search engine is based on neural network theory and statistics to quickly analyze massive information, and quickly classify it, and transmit the information and data related to the input search content to the user. In addition, the intelligent search engine can also automatically grasp the complex concepts in the new information data, automatically learn the meaning of the concepts, and then store the new concepts, which is convenient for searching and searching for the concepts in the future [2]. Adding neural network and pattern recognition to intelligent search will enable intelligent search to locate and search for information based on the text content, contextual connection, and meaning. Therefore, the

intelligent search can quickly find the content which is the closest and most useful information, and it can also provide users with an expandable traditional system of information, data, and knowledge, thus greatly improving the utilization rate of teaching resource intelligence. The system framework of intelligent information retrieval is shown in Figure 3.

This model is mainly composed of two relatively independent parts: an offline processing part and an online processing part. The offline processing part is mainly responsible for the collection of original information, the establishment of domain ontology, and the processing of original information. The online part of the processing mainly includes obtaining the user's query request, query condition coding, querying the semantic index library through the search engine, sorting the retrieved result set, and returning the sorted final result to the user. The offline processing part is relatively time-consuming, so you can choose to disconnect the service and the server load is light.

3. Design of Educational Resource Search Based on Artificial Intelligence Search Algorithm

3.1. Characteristics of Intelligent Search Engine in Teaching Platform. Traditional search engines have many problems in function, such as a large amount of information feedback in search results and more irrelevant information. This is mainly because traditional search engines mainly use keyword-based mechanical matching methods to achieve users' search needs [21]. To a certain extent, the ability to understand and analyze search sentences is not accurate enough; that is to say, traditional search engines cannot perform fast and accurate searches for users' personalized knowledge and professional knowledge in different fields. However, with the further development of search engine technology, the emergence of smart search engines can solve the above problems [7]. Smart search engines can provide users with a friendly search interaction interface and can improve the original resource search based on keyword search to knowledge understanding. They can search for the most meaningful resources for users from existing resources, give feedback according to the user's search request, and can realize automatic word segmentation technology, synonym technology, machine learning technology, etc. Intelligent search engines will have more personalized features, can gradually understand users' search preferences based on information such as resource browsing habits and search habits of different users, establish corresponding user description model information for different users, and actively provide users with relevant information, to bring users a better search experience [22].

The in-site intelligent search engine framework proposed in this article is designed for domestic learners of the teaching platform. Through the investigation of the use of the teaching resource library in the online teaching platform, it is found that most of the resources in the resource library are documents, mainly including document resources such as teaching courseware and extracurricular reference

materials in such formats. Therefore, the intelligent search system in the teaching platform should mainly have the following five characteristics:

- (1) Intelligent and personalized search function: the intelligence of search engines is mainly reflected in the ability to obtain users' search preferences according to their search habits and then provides different users with different types of information. In this article, the search engine should be able to construct a learner's interest model according to the learner's characteristics and update this interest model in real time, so as to determine the scope of the search and provide its professional and personalized search service
- (2) Facing the field of teaching resources and strong professionalism: at present, general-purpose commercial search engines cannot well meet the personalized search needs of users in a certain field. The in-site intelligent search engine in the teaching platform realizes the function of quickly searching for the teaching resources in the platform and obtains the corresponding teaching resources according to the theme, which shortens the time for learners to obtain resources
- (3) Good user interface: when designing the user interface, you can refer to the current popular search engines, combine the functions of the online teaching platform itself, and add in-site links as appropriate. For example, you can specify which courses the teaching resources come from, and which specific teaching modules in the course, etc. improve the user's search experience.
- (4) Providing learners with search recommendations: the on-site teaching resource search engine is a search engine oriented to a specific field. The search tool needs to provide learners with some more commonly used keyword input prompts related to teaching resources, use natural language to interact with learners, and adopt technologies such as semantic web and automatic word segmentation technology which accurately and quickly understands learners' search requests
- (5) Higher search result accuracy rate: search engines use information filtering, information mining, and other technologies to filter the teaching resources in the network teaching platform and combine the search sentences to filter the importance of learners again, so that the feedback search results are more in line with the learners' search needs

3.2. Artificial Intelligence Applied to Search Engine Resource Acquisition. From the above analysis, we can know that the existing search engine resource acquisition methods are blind. Relying on width and depth-first algorithms, they often get a lot of irrelevant information, which leads to the decline of efficiency and search accuracy. So how can we improve the efficiency and accuracy of search and how can

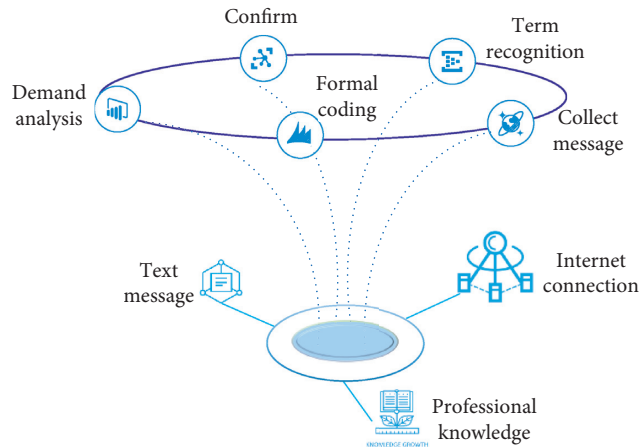


FIGURE 1: Prototype method of domain ontology construction.

we obtain more valuable information resources at the same time? If the required information is on the page and the depth-first search algorithm is used, then the previous must be traversed; it is the turn to visit the page after a page, the price paid is very large, but the efficiency is only about 10%. If the heuristic search is used, generally speaking, it can be accessed as long as the page is traversed, which is much more efficient than the depth-first algorithm. The application of artificial intelligence to web crawlers will enable search engines to achieve greater success in the acquisition of information resources [23]. Using heuristic algorithms, web crawlers can eliminate irrelevant links when traversing links, and only visit and traverse links that are deemed useful after the evaluation function are calculated. This greatly reduces the number of traversed links, and the obtained pages that meet the requirements account for a large proportion of the entire traversed pages. The accuracy of the collection of information resources by web crawlers has also improved. Intelligent agents, also known as agents, are new results of artificial intelligence research. It can perform various complicated tasks on behalf of the user according to the user's needs without clear specific requirements. The application of intelligent agent technology to search engines will provide a broad space for the development of a new generation of more powerful online intelligent search systems. When accessing a web page link, how to judge the type of the page is a tricky problem. If you can develop an intelligent agent system, let it automatically complete the classification of documents or judge its type; it would be a better way. The main processing steps taken by the artificial intelligence search algorithm are shown in Figure 4.

The so-called network information retrieval also refers to Internet retrieval. With the function of network interface software, users can query information in various fields on the terminal. This retrieval method is mainly based on the distribution function of the Internet, that is, data distribution and storage, and a large amount of information is distributed in the corresponding server. The user searches for the dispersedly stored data through the application terminal system, and any data can be searched and applied

on the network. The basic structure of the network information retrieval system is shown in Figure 5.

First, take the knowledge base as an example. As the part responsible for long-term memory in the artificial intelligence system, this part mainly stores specialized knowledge. This includes established facts and information, as well as general common sense and rules. Some special systems also contain databases. The second is the inference engine. As a set of programs for performing retrieval tasks, it mainly includes master control and various task programs. As a specialized library, it plays an important role in providing search support. The user interface mainly includes the system and the links used by the user to input and convey relevant information. As a bridge between external information and internalization, it can display the final processing result to the user and transfer the user's will to the computer. The nonnatural language used in this process can also reduce the burden and pressure of users in the process of use. The intermediate database is also the blackboard we often say; it is mainly used to store intermediate results and data in the process of performing tasks and inferences. In the actual application process, the system will first display the problem on the blackboard and then display the initial state of the problem. Then, the expert system makes inferences based on the knowledge search matching status of the knowledge base and, at the same time, constantly repairs and infers the contents of the blackboard. When necessary, we will also ask customers to make up for and solve the shortage of knowledge in the knowledge base. Therefore, the blackboard can also be regarded as a dynamic knowledge base to a certain extent. It plays a very critical role in controlling the progress of the job. Finally, the knowledge acquisition device is also called a learner. Its main function is to continuously repair the content of the knowledge base based on the experience of system operation. This is an effective supplement to ensure the smooth operation of the system. The interpreter is mainly responsible for solving the user's questions and is also responsible for explaining to the user the calculation trajectory of the system conclusion. In general, artificial intelligence technology is a means of science and

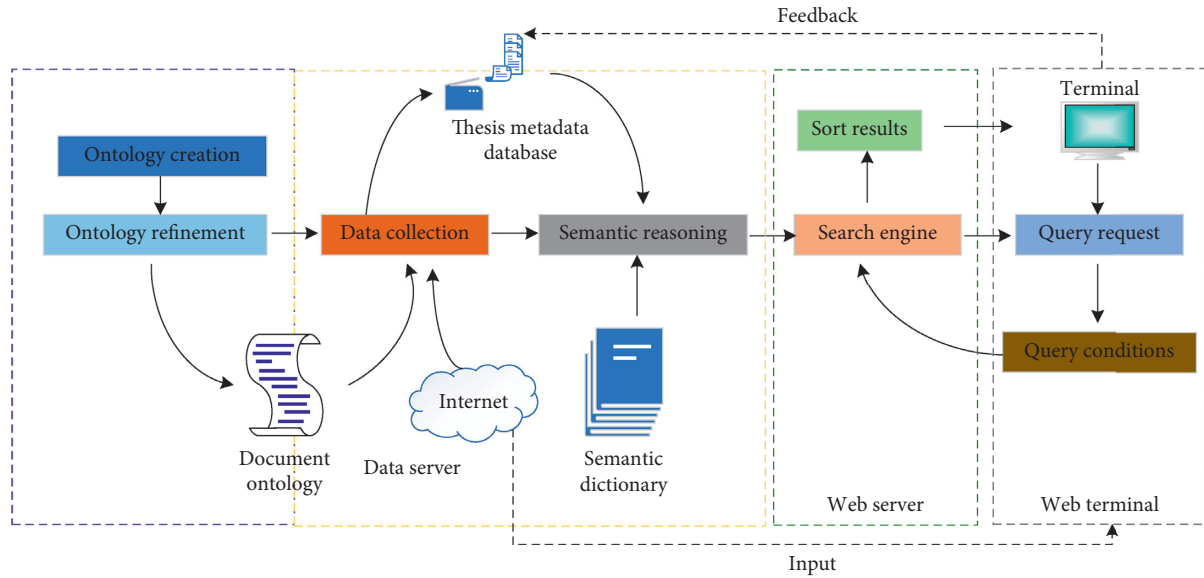


FIGURE 2: The system structure of intelligent search.

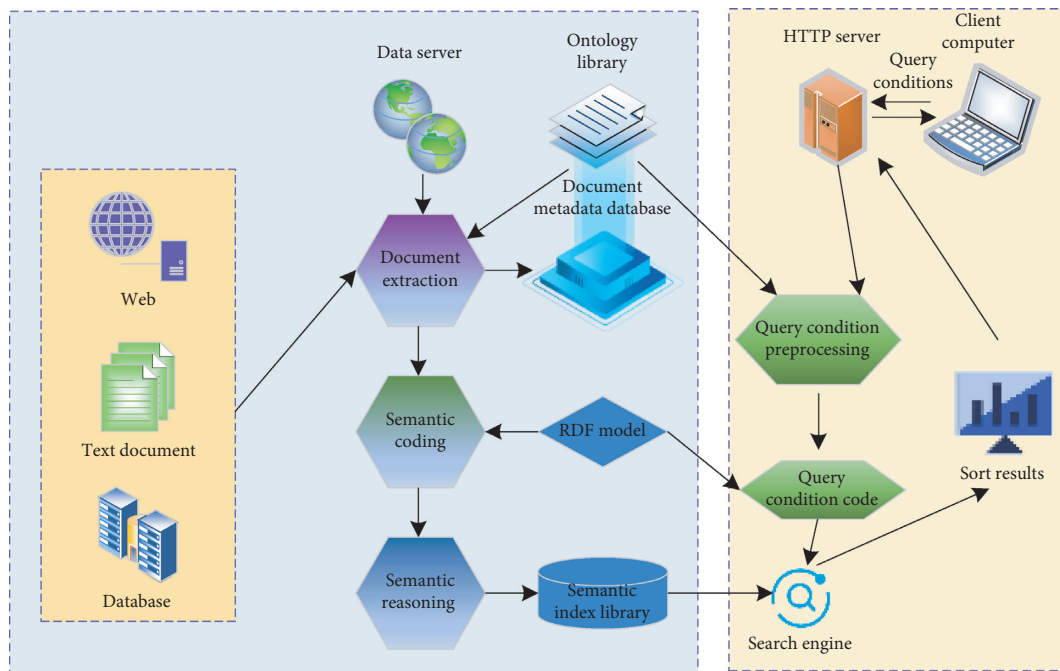


FIGURE 3: Framework for intelligent information retrieval.

technology to help people solve some urgent problems in the form of analogy thinking.

4. Experimental Results and Analysis

4.1. Data Set and Experiment Establishment. In order to objectively evaluate the abovementioned feature selection method in the actual situation of basic education text classification, we select basic education discipline resources to construct a hierarchical classification system with a two-tier structure. As shown in Figure 6, there are 12 major

categories and 18 subcategories and manually obtained a corresponding data set, which includes 8428 training documents and 1867 open test documents. The specific distribution is shown in Figure 6. In the experiments and discussions later in this article, unless otherwise specified, they refer to the training set and test set.

The classification algorithm used in the experiment is the Roccio algorithm. In the actual application, $\alpha = 1$ and $\beta = 0$ in the Roccio formula, that is, only positive samples are considered, and the category vector is the center of the training text of the category, because in many cases, the effect of

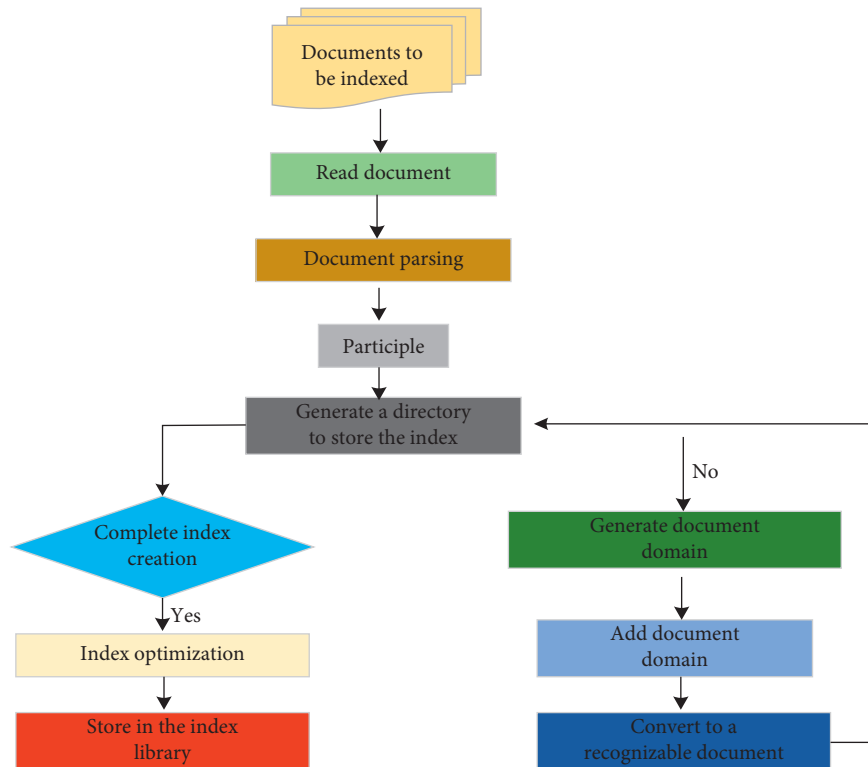


FIGURE 4: The process of index creation.

negative samples on the establishment of the category vector is limited, and it may even bring more noise. The classification results are evaluated using the microaverage classification accuracy rate Micro-F and the value evaluation standard. The method of composing training documents is as follows: the training documents of the first-level category are composed of documents in the subcategories below them. For the first-level categories, when training a certain category, the positive example document set is all the documents of that category, and the negative example document set is the sum of the documents of other categories. The second level of the subcategory training method is to compare each subcategory under the same category. For example, for the subcategories of the language category elementary school language, junior high school language, and high school language, when extracting the characteristic words of the elementary school language category, the normal document set is a document of elementary school language, and the negative example document set is the sum of documents of junior high school language and high school language. This improves the training speed, and at the same time, the feature distinction between subcategories is greater. In the test, a top-down classification method is adopted; that is, the document is first classified into a certain category, and then, the document is then classified by each small category under the category to continue to determine whether it belongs to a small category.

4.2. Analysis of Correct Rate Results. The first round of experiments compared six commonly used feature selection

methods of DF, IG, Cross Entry, CHI, WET, and MI. Go to the second layer, which is the Micro-F of the subcategory, and the value is shown in Figure 7. It can be seen from the two figures that, in addition to mutual information, other algorithms have achieved better classification results. Among them, the Micro-F of the major category has a very high value, with the maximum value reaching 97.1%. With the decrease in the number of features, the overall trend is declining, but the Micro-F values of various methods are not much different before 3000 dimensions. The Micro-F value of text evidence weight is significantly reduced after 300D dimensions. Another Micro-F value shows an obvious downward trend after 1000 dimensions. The Micro-F value of the small category is up to 78.2%, which is much lower than the Micro-F of the large category. This may be because the discrimination between the small categories under the same category is relatively poor, resulting in poor classification results, but the overall trend graph is similar to that of major categories.

In feature selection, the function of feature in classification is expressed as the weight to select important features. In the text description, we need to give weight to the features to describe the importance of the text features in the text. In essence, the consideration of feature importance in these two places is the same, which is based on the classification ability and description ability of features. However, the scope of their consideration is different, and the methods adopted are also different. In the face of the whole data set, the importance of features is for all categories, and the distribution of features among categories is emphasized. The latter is for a single text, and the importance of features is for the content

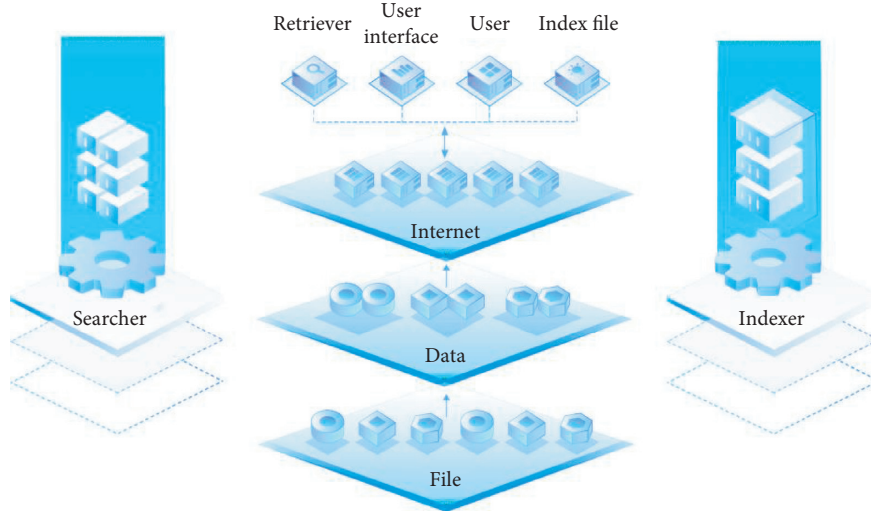


FIGURE 5: The basic process of intelligent information retrieval.

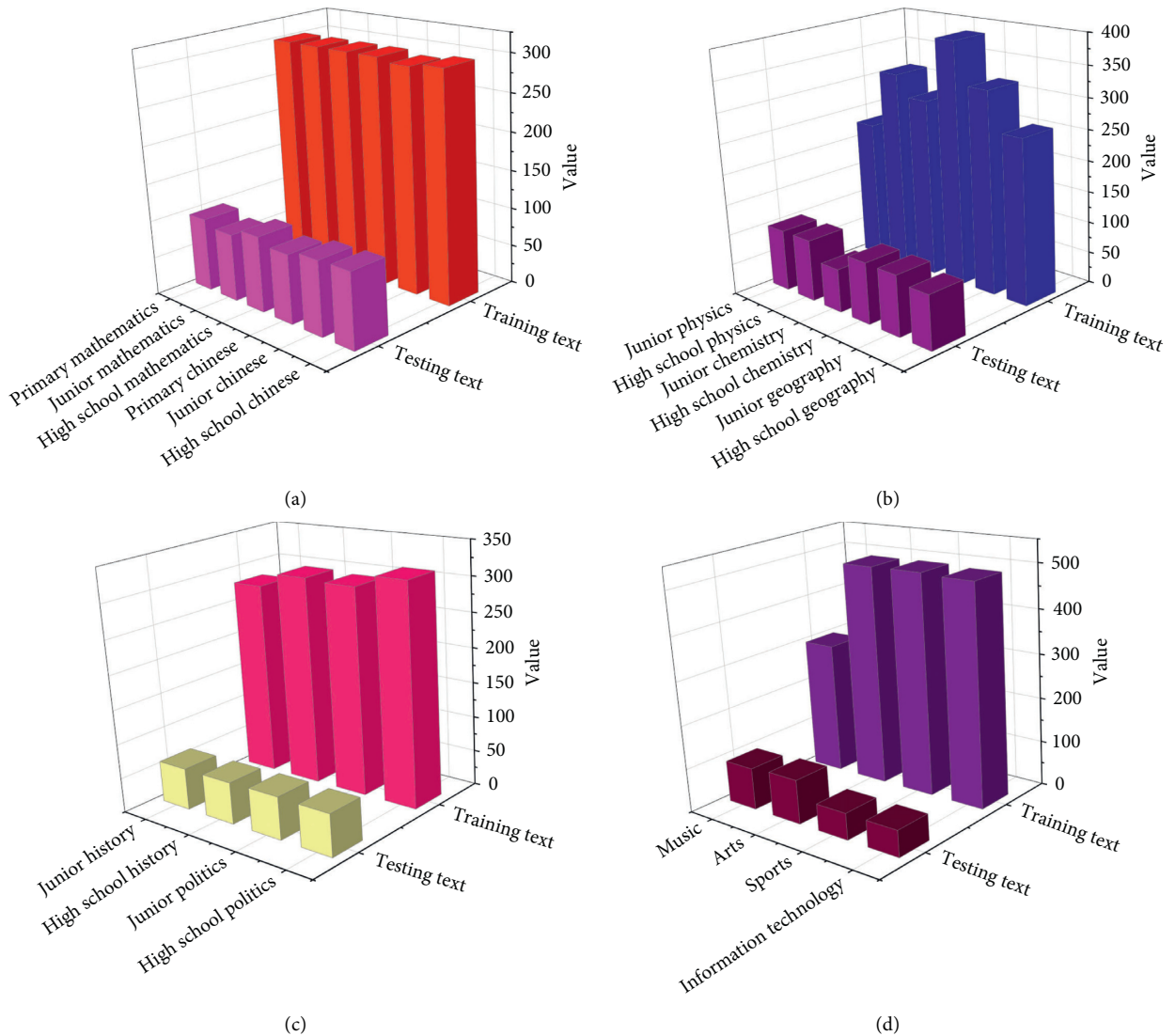


FIGURE 6: Distribution table of the training set and test set.

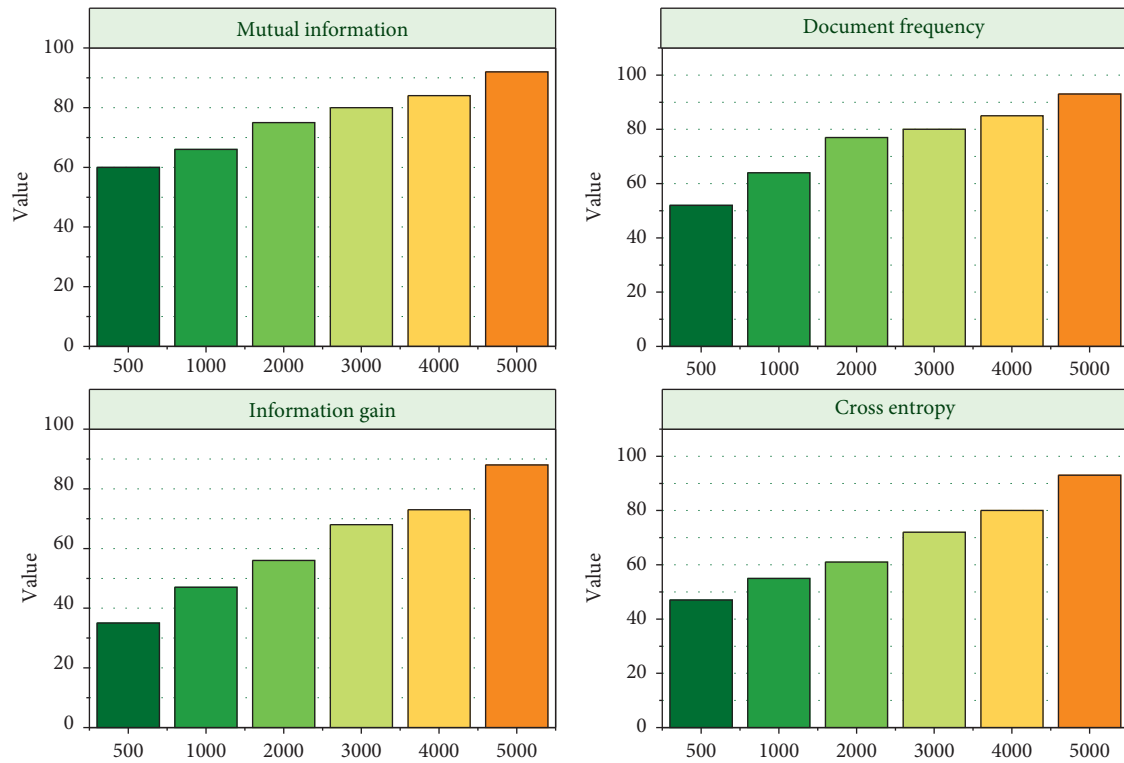


FIGURE 7: The accuracy of the first kind of microaverage classification.

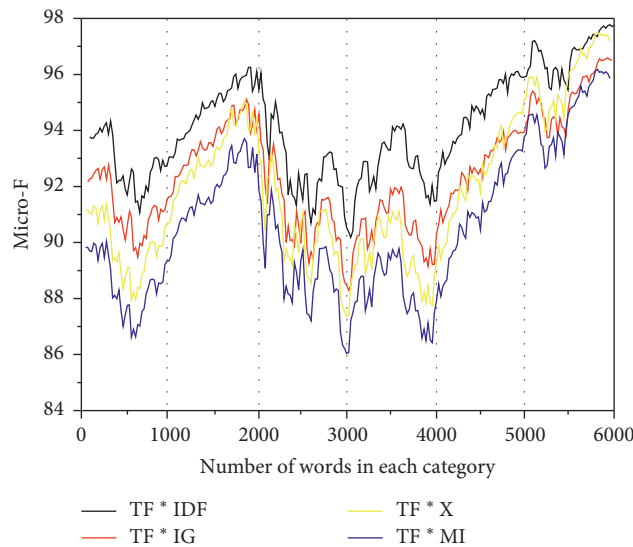


FIGURE 8: Accuracy of microaverage classification.

of the text. In order to measure the description ability and classification ability of features in the text, it also needs to be placed in the global scope to measure accurately.

4.3. Analysis of Classification Results. The Roccio classifier is used to perform weighting experiments on 1867 documents in 12 categories and 18 subcategories of the basic education text set. In this experiment, no dimensionality reduction

algorithm is used. The microaverage Micro-F value test is used for the test results. Figure 8 shows the accuracy of the microaverage classification of texts classified into major categories. It can be seen from the figure that TF*IDF has the best classification effect, TF*IG and TF*X are slightly worse than TF*IDF, and TF*MI has the worst effect, followed by TF*WET and TF*MIPW. The overall trend of TF*IDF is that as the feature dimension decreases, the classification effect becomes worse, but at 3000 dimensions,

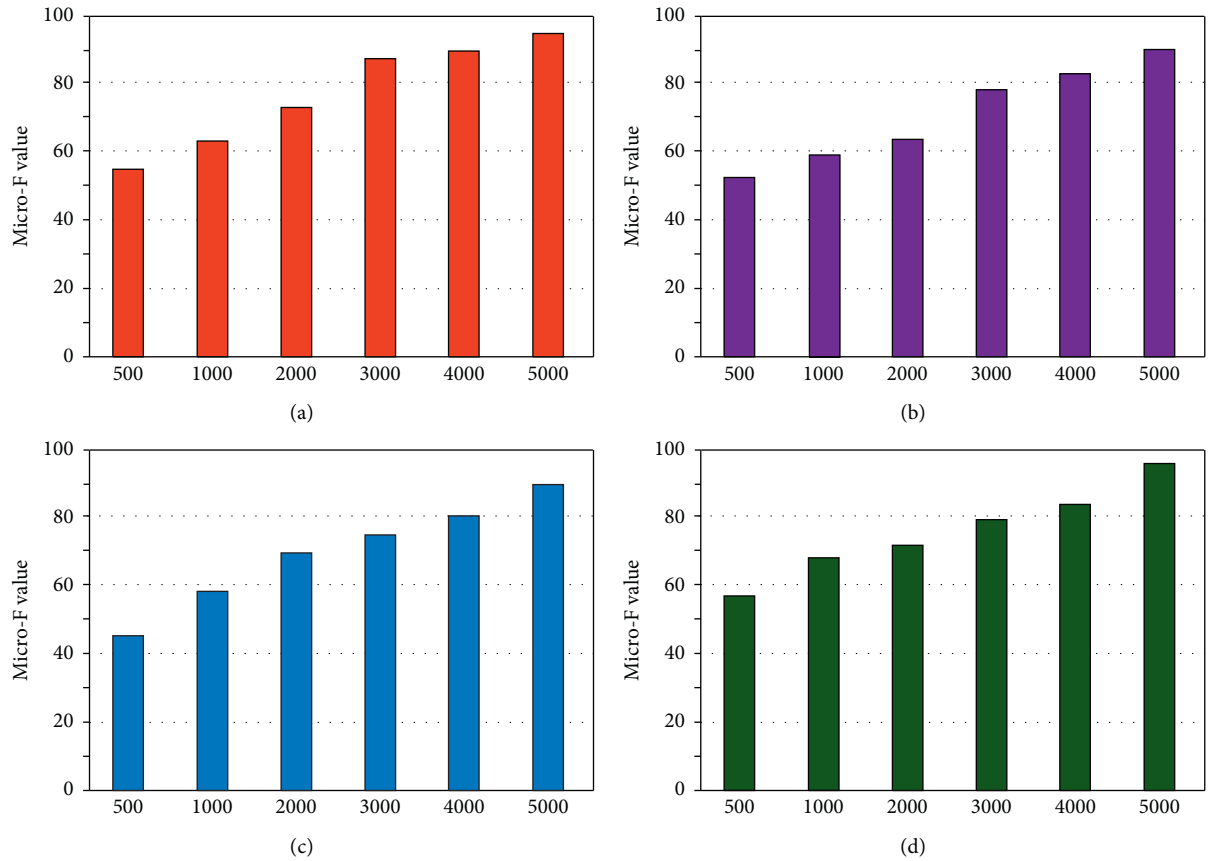


FIGURE 9: Accuracy of microaverage classification. (a) Classification results of TF * IDF. (b) Classification results of TF * IG. (c) Classification results of TF * X. (d) Classification results of TF * MI.

the classification accuracy rate slightly increases. Among them, the Micro-F value is 0.964.

Figure 9 shows the accuracy of the microaverage classification of texts classified into subcategories. It can be seen from the figure that TF*IDF has the best classification effect, TF*Igorot, TF*X', and TF*IG are slightly worse than TF*IDF, and TF*MI has the worst effect, followed by TF*WET and TF*MIPW. Except for TF*MI, the overall trend of other functions is that as the feature dimension decreases, the classification effect becomes worse, but the decline is small before 3000 dimensions, and the decline is greater after 3000 dimensions. The Micro-F values of TF*IDF, TF*Igorot, and TF*X² are equivalent at 500 dimensions. At 3000 dimensions, the Micro-F value of TF*IDF is 0. It can be seen that, compared with the classification effect of the large class, the classification accuracy of the small class is still much worse. This may be because, compared with the features of the large class, some small characteristics of the class are not clearly distinguished because of the reasons.

5. Conclusion

With the continuous growth of information on the Internet, people have become a way of learning to find what they need

through the Internet. In particular, for students in education, their learning is based on resource learning. Therefore, obtaining useful subject knowledge is a key issue for online education students to improve their knowledge. However, due to the limitations of students' knowledge and the control of the Internet, it is difficult to obtain the required knowledge from a general search. Therefore, the development of intelligent search is now increasingly becoming the focus of search research in teaching. With the development of artificial intelligence technology, the intelligence of search engines will become more and more in line with people's actual needs. This article is based on the knowledge in the field of artificial intelligence as the background; according to the characteristics of teaching resources, it is beneficial to search for teaching resources, explore, and build a semantic network in its retrieval module, aiming to classify the domain words in educational resources, so that the retrieval engine can expand the query request based on this concept. This article is to study the use of intelligent search in teaching resources. It is a preliminary exploration of intelligent search engines. Although some methods are proposed, there are still many problems to be studied. This article uses the word frequency method to calculate the degree of correlation between the domain words and the document. It is more practical under certain conditions, but it cannot objectively

represent the content of the document. Therefore, how to understand and express the content of the document has always been the intelligence of the search engine. The important direction of transformation is not only to be able to extract the main keywords from the document but also to understand the main content of the document, to automatically locate the document accurately, and truly realize the understanding of the conceptual level of the document. With the development of artificial intelligence technology, the key technology of intelligent search will inevitably be greatly improved, and the intelligent search will certainly have broad application prospects in the search field in the future.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] S. Li and Y. Wang, "Research on interdisciplinary characteristics: a case study in the field of artificial intelligence," *IOP Conference Series Materials Science and Engineering*, vol. 67, no. 7, pp. 52–53, 2019.
- [2] S. Huang, "Design and development of educational robot teaching resources using artificial intelligence technology," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 5, pp. 116–118, 2021.
- [3] I. S. Kang, J. W. Moon, and J. C. Park, "Recent research trends of artificial intelligent machine learning in architectural field-review of domestic and international journal papers-," *Journal of the Architectural Institute of Korea Structure & Construction*, vol. 33, no. 4, pp. 63–68, 2017.
- [4] J. W. Liang and D. B. Wang, "Analysis of research progress in the field of artificial intelligence in the past ten years," *Jiangsu Science and Technology Information*, vol. 36, no. 11, pp. 22–27, 2019.
- [5] Y. Li, "Research on the construction of TCFL resource database system based on artificial intelligence," *Journal of Intelligent and Fuzzy Systems*, vol. 11, no. 6, pp. 1–12, 2019.
- [6] C. Lu, M. J. Guo, F. Y. Zhang et al., "Stakeholder analysis of the application of artificial intelligence in the field of primary health care," *Acta Academia Medicine Sinical*, vol. 43, no. 1, pp. 101–108, 2021.
- [7] H. Liang, "Role of artificial intelligence algorithm for taekwondo teaching effect evaluation model," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 3239–3250, 2021.
- [8] L. Feng, J. Wang, C. Ding et al., "Research on the feedback system of face recognition based on artificial intelligence applied to intelligent chip," *Journal of Physics: Conference Series*, vol. 17, no. 3, pp. 32–34, 2021.
- [9] J. Liu, T. Li, J. Chen, and F. Zuo, "Research on improved intelligent control processes based on three kinds of artificial intelligence," *Processes*, vol. 8, no. 9, pp. 1042–1043, 2020.
- [10] G. T. Vu, B. X. Tran, R. S. McIntyre et al., "Modeling the research landscapes of artificial intelligence applications in diabetes (GAPRESEARCH)," *International Journal of Environmental Research and Public Health*, vol. 17, no. 6, pp. 1982–1985, 2020.
- [11] J. Hu, "Teaching evaluation system by use of machine learning and artificial intelligence methods," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 5, pp. 87–89, 2021.
- [12] A. M. Cox, S. Penfield, and S. Rutter, "The intelligent library: thought leaders' views on the likely impact of artificial intelligence on academic libraries," *Library Hi Tech*, vol. 37, no. 5, pp. 36–38, 2018.
- [13] T. Liang, "Exploration and application of the "Small-Class discussion" teaching mode in the field practice teaching of general geology," *Creative Education Studies*, vol. 7, no. 5, pp. 623–627, 2019.
- [14] Y. Chen, "Research on the development of Chinese intelligent toy industry in the artificial intelligence age," *Science Technology and Industry*, vol. 17, no. 10, pp. 15–18, 2017.
- [15] M. Chen and M. J. Zhang, "Evaluation and research on the design of intelligent classroom teaching in primary school Chinese," *Education Teaching Forum*, vol. 10, no. 7, pp. 235–239, 2019.
- [16] X. Li, "The construction of intelligent English teaching model based on artificial intelligence," *International Journal of Emerging Technologies in Learning*, vol. 12, no. 2, pp. 35–44, 2017.
- [17] B. Hu and B. C. Li, "Applications of artificial intelligence in intelligent manufacturing: a review," *Frontiers of Information Technology & Electronic Engineering*, vol. 18, no. 1, pp. 86–96, 2017.
- [18] M. Gao, J. Xing, C. Yin et al., "Personalized recommendation method for English teaching resources based on artificial intelligence technology," *Journal of Physics Conference Series*, vol. 17, no. 1, pp. 12–14, 2021.
- [19] B. Garlic, "The application of artificial intelligence in the process of optimizing energy consumption in intelligent areas," *Neural Network World*, vol. 27, no. 4, pp. 415–446, 2017.
- [20] C. Y. Chang, "How the intelligent recognition industry service (IRIS) centre is driving technological research in artificial intelligence," *Impact*, vol. 20, no. 11, pp. 12–14, 2021.
- [21] C. Yang, S. Huan, and Y. Yang, "A practical teaching mode for colleges supported by artificial intelligence," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 15, no. 17, pp. 195–197, 2020.
- [22] M. J. Timms, "Letting artificial intelligence in education out of the box: educational cobots and smart classrooms," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 701–712, 2016.
- [23] R. Duan, Y. Wang, and H. Qin, "Artificial intelligence speech recognition model for correcting spoken English teaching," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 1, pp. 1–12, 2020.

Research Article

A Temporal Pool Learning Algorithm Based on Location Awareness

Lei Li , Yuquan Zhu , Tao Cai , Dejiao Niu , Huaji Shi , and Tingting Zou 

Department of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China

Correspondence should be addressed to Lei Li; lilei@ujs.edu.cn, Tao Cai; caitao@ujs.edu.cn, and Dejiao Niu; djniu@ujs.edu.cn

Received 22 March 2021; Revised 29 April 2021; Accepted 2 June 2021; Published 11 June 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Lei Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hierarchical Temporal Memory is a new type of artificial neural network model, which imitates the structure and information processing flow of the human brain. Hierarchical Temporal Memory has strong adaptability and fast learning ability and becomes a hot spot in current research. Hierarchical Temporal Memory obtains and saves the temporal characteristics of input sequences by the temporal pool learning algorithm. However, the current algorithm has some problems such as low learning efficiency and poor learning effect when learning time series data. In this paper, a temporal pool learning algorithm based on location awareness is proposed. The cell selection rules based on location awareness and the dendritic updating rules based on adjacent inputs are designed to improve the learning efficiency and effect of the algorithm. Through the algorithm prototype, three different datasets are used to test and analyze the algorithm performance. The experimental results verify that the algorithm can quickly obtain the complete characteristics of the input sequence. No matter whether there are similar segments in the sequence, the proposed algorithm has higher prediction recall and precision than the existing algorithms.

1. Introduction

Hierarchical Temporal Memory (HTM) is a new neural network model that simulates the organization and structure of cortical cells and the way the human brain processes information. HTM uses a large amount of sequence data for training [1]. By extracting and storing the patterns, HTM can predict the subsequent occurrence of the current input or detect whether the current input is abnormal. Therefore, HTM is widely applied in the analysis and processing of time series data. The basic unit of HTM is the cell. A number of cells constitute the minicolumn, and a large number of minicolumns form the network space of HTM. Spatial pool learning algorithm (SPL) and temporal pool learning algorithm (TPL) are two important learning steps of HTM. SPL establishes the mapping relationship between the input and the minicolumn set to form the sparse distributed representation (SDR) of inputs. TPL constructs or adjusts the corresponding dendrite branches among the cells to form a temporal correlation of inputs [2].

In the process of human learning, the human brain remembers many things. When given clues, the human brain can retrieve relevant knowledge with high efficiency. HTM works by imitating the way the human brain processes information. Using the memorized patterns, HTM can also efficiently obtain the complete pattern information by the clues. With SPL, the input can be quickly located and represented in HTM. TPL uses different combinations of active cells to express the context of input and constructs the temporal relationship between inputs in different environments. When an input enters HTM, HTM can quickly retrieve all follow-ups by activating all cells on the active minicolumn. With the continuous confirmation of the following inputs, HTM can identify and confirm the patterns containing sequential inputs. The retrieval efficiency of this method is much higher than that of the traditional comparison algorithm. Based on this, the trained HTM can be regarded as a warehouse of rules, which can detect whether the test sequence is abnormal. In this model, if the learning sequences are the rules, it is

worth discussing how to get all features of the rules quickly and ensure that the trained HTM has high prediction recall and precision.

First, TPL uses Hebbian rules to adjust the dendritic branches of cells to establish the temporal association between adjacent inputs. When the number of connected synapses in the dendrite branch of the cell accumulates to a certain threshold, TPL completes the extraction and memory of sequence features. TPL needs multiple iterations to memorize the sequence features, which affects the learning efficiency of TPL. Secondly, TPL gives priority to enhance the sequence features that have been learned rather than the current position sequence features, which affects the learning effect of TPL. In particular, when there are many continuous same inputs in the sequence, TPL is easy to fall into the same input cycle training and unable to obtain all features of the pattern, which affects the learning effect of TPL. Therefore, it is necessary to design a new algorithm to improve the learning efficiency and effect of TPL.

In this paper, a temporal pool learning algorithm based on location awareness (TPL_LA) is proposed. The generation rules of learning cells and active cells based on location awareness and the adjustment rules of dendrite branches associated with adjacent inputs are designed. Three datasets are used to test the performance of TPL_LA and TPL. The main innovations of this paper are as follows.

- (1) Using the multiple cells contained in the minicolumn, the generated learning cells and active cells can accurately express the position information of the input. In this way, the algorithm can give priority to learning the current sequence features, reduce the possibility of self-cycle prediction when learning continuous same inputs, and improve the learning effect of TPL.
- (2) According to the generated learning cell set and active cell set, the algorithm adjusts the updating rules of dendritic branches associated with adjacent inputs, and a new setting strategy of synaptic value is designed to improve the learning efficiency of TPL.
- (3) According to the generated learning cell set, the algorithm reduces the scope of active cells to predict the follow-up more accurately, which reduces the impact of historical memory on learning
- (4) In the open-source framework of HTM, the prototype of TPL_LA is implemented, and different types of data sets are used to test and analyze it. The results verify that TPL_LA could effectively improve the learning efficiency and effect compared with TPL.

The remainder of this paper is structured as follows: the related works for HTM are discussed in Section 2. Then, in Section 3, the efficient temporal pool learning algorithm based on location awareness is discussed. Finally, the experiments and analysis are carried out in Section 4. And the last section concludes our study.

2. Related Works

HTM is a new artificial neural network model based on Jeff Hawkins' memory prediction framework [3], and its main workflow is shown in Figure 1. The basic network space of HTM can be described as a region composed of multiple minicolumns. Regions and regions can have hierarchical relationships. There are many cells on a minicolumn. The cells on each minicolumn share a proximal dendrite to receive input stimuli. Every cell may have many branches of distal dendrites, which can build relationships with other cells [1]. The encoder converts the input into a binary vector of equal length, and the number of components with 1 in the vector is equal.

The vector can activate a small number of minicolumns to express the input content through SPL. TPL selects some cells on these activated minicolumns to express the input location information and establishes dendrite branches on these cells to construct the correlation between inputs. The sequence can be learned through the SPL and TPL to form memory in the network space of HTM. When the input sequence is consistent with the memorized pattern, the HTM can predict the next content based on the memory [4]. Two core algorithms, SPL and TPL, are used to construct the space-invariant and time-invariant features of input sequences, which are the key technologies for HTM to distinguish similar patterns.

For SPL, Ahmad and Hawkins give the characteristic description of sparse matrix and the expression form of information in the spatial pool [5], which provides theoretical guidance for the following research. Lattner Stefan uses the vector to study the spatial pool and proposes a formula for calculating the overlapping values and the corresponding learning rules [6]. Byrne describes the structure of HTM formally by using matrixes [7], which provides a mathematical basis for the derivation of rules. Leake et al. study and discuss the influence of initialization parameters on SPL [8]. Mnatzaganian et al. give a comprehensive mathematical framework of the spatial pool [9] so that the follow-up scholars can understand the core characteristics of SPL and improve the algorithm from a mathematical point of view. Li et al. propose an improved algorithm for spatial pool learning, which can make the training process of the spatial pool more efficient and stable [10]. In recent years, increasing scholars have begun to study the hardware implementation of the spatial pool to give full play to the advantages of HTM architecture [11]. Deliang Fan et al. propose that the computing blocks for HTM can be mapped using low-voltage, magneto-metallic spin-neurons combined with an emerging resistive crossbar network [12]. Krestinskaya et al. develop circuits and systems to achieve the optimized design of an HTM SP, an HTM TM, and a memristive analog pattern matcher for pattern recognition applications [13]. SPL abstracts the input features through a hierarchical structure [14], which makes HTM have wide-ranging applications in recognition and classification, such as data classification [15], face recognition [16], speech recognition [17], biometric recognition [18], detection of multiple objects located in clutter color images [19],

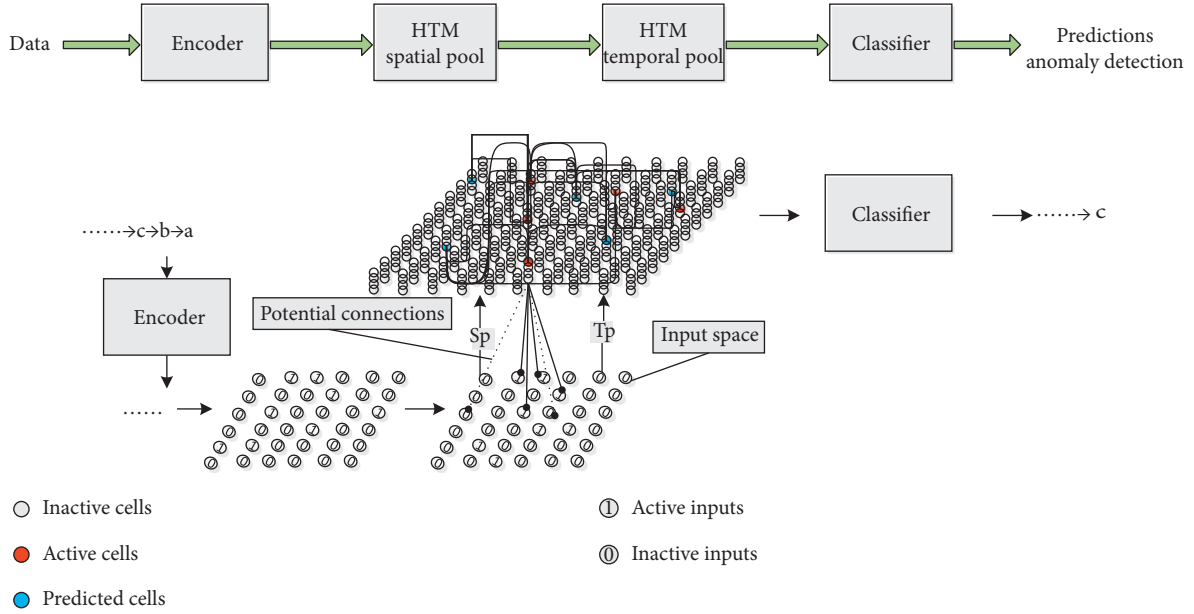


FIGURE 1: HTM workflow.

handwriting recognition [20], action recognition [21], gait recognition and understanding [22], and natural language processing [23].

For TPL, Hawkins and Ahmad give the formal expression of learning rules by matrixes [2], which gives a mathematical description of the rules related to active cells, learning cells, predicting cells, and synaptic renewal. This provides a theoretical basis for the subsequent study of the temporal pool learning algorithm. Younes Bouhadjar et al. reformulate the model in terms of a network of spiking neurons with continuous-time dynamics [24], in which they discuss how neuronal parameters such as cell-intrinsic time constants and synaptic weights constrain the sequence-processing speed. El-Ganainy et al. propose an improved Hierarchical Temporal Memory that can consider long-term dependence [25]. Suzugamine et al. propose a self-structured cortical learning algorithm that dynamically adjusts the numbers of columns and cells according to the input data [26]. In this model, the backtracking algorithm proposed by the author makes the prediction of the temporal pool more stable. In recent years, some scholars start to implement TPL on hardware [27]. Many scholars combine SPL with TPL and apply HTM to many fields with time series data, such as anomaly detection of time series data [28], heart attack detection [29], medical data flow prediction [25], hydrological intelligent monitoring [30], abnormal ECG detection [31], and abnormal detection in crowd management [32]. However, when the sequence data is learned as rules, TPL in the above research still has some limitations. For example, it needs multiple iterations to memorize the sequence features, it gives priority to enhancing the memory of a historical pattern when learning a sequence, and it is easy to fall into the same input cycle training when learning the sequence with many continuous same inputs. When TPL obtains all features of sequences, these limitations will lead to low learning efficiency and poor learning effect.

3. The Temporal Pool Learning Algorithm Based on Location Awareness

In TPL_LA, the learning process of TPL is reconstructed. The active cell and learning cell selection rules based on location awareness are proposed to make TPL_LA more focused on the current sequence features, which improves the learning effect of the algorithm. The dendrite updating rules associated with adjacent inputs are proposed to improve the learning efficiency of TPL_LA. The algorithm reduces the scope of active cells to predict the follow-up more accurately, which reduces the impact of historical memory on learning. The learning process of TPL_LA is shown in Figure 2.

3.1. Notation. The following notations are given for the HTM consisting of N minicolumns and M cells in each minicolumn:

- (1) W^t : the set of active minicolumns in HTM at time t . The set includes the identifiers of the activated minicolumns at time t . W^t is the output of SPL.
- (2) $A^t = \{a_{ij}^t\}$: the matrix of active cells at time t . A^t is a binary matrix of $M \times N$. a_{ij}^t indicates whether the i -th cell on the j -th minicolumn is active at the time of t . A value of 1 indicates an active state and a value of 0 indicates an inactive state.
- (3) $\Pi^t = \{\pi_{ij}^t\}$: the matrix of predictive cells at time t . Π^t is a binary matrix of $M \times N$. π_{ij}^t indicates whether the i -th cell on the j -th minicolumn is predictive at the time of t . A value of 1 indicates a predicted state and a value of 0 indicates an unpredicted state.
- (4) D_{ij}^k : the matrix of dendrite branch synaptic. D_{ij}^k is a binary matrix of $M \times N$. j is the minicolumn's number. i is the cell number on the minicolumn. k is the dendrite branch number of the cell. The matrix is

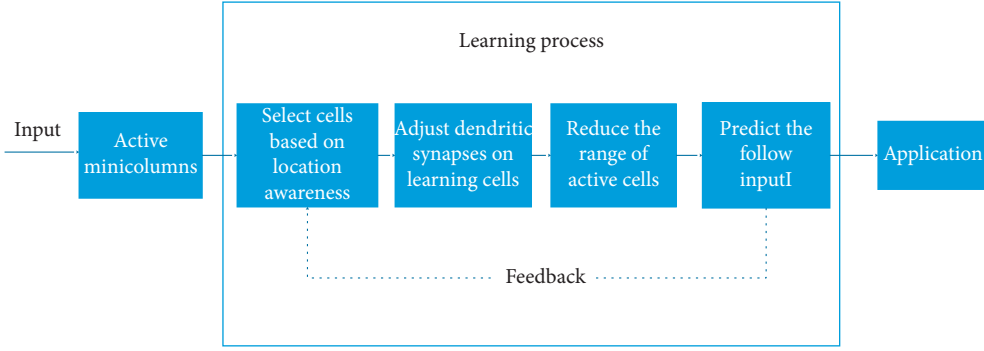


FIGURE 2: The learning process of TPL_LA.

used to store the synaptic permanence in the k -th dendrite branch between i -th cell of j -th minicolumn and other $m \times n^{-1}$ cells. Permanence is a scalar value, and its range is $[0.0, 1.0]$.

- (5) \tilde{D}_{ij}^k : the matrix of dendrite branch connectivity. \tilde{D}_{ij}^k is a binary matrix of $M \times N$ constructed by the following equation. C_P is the threshold of connectivity.

$$\tilde{d}_{ij}^k = \begin{cases} 1, & d_{ij}^k \geq C_P, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

- (6) \dot{D}_{ij}^k : the matrix of positive terms of dendrite branch synapses. \dot{D}_{ij}^k is a binary matrix of $M \times N$ constructed by the following equation:

$$\dot{d}_{ij}^k = \begin{cases} 1, & d_{ij}^k > 0, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

- (7) $\dot{\text{Cell}}^t = \{\dot{\text{cell}}_{ij}^t\}$: the matrix of learning cells at time t . $\dot{\text{Cell}}^t$ is a binary matrix of $M \times N$ which is a subset of A^t . $\dot{\text{Cell}}^t$ contains the cells whose dendrite branches need to be added or modified in the training process at time t . When $\dot{\text{cell}}_{ij}^t = 1$, it means that the cell participates in the learning process of the current input.

- (8) I_P : initial synaptic permanence. I_P is a scalar with a value range of $[0.0, 1.0]$, indicating the permanence of the first connection between two cells.

3.2. Analysis of the Temporal Pool Learning Algorithm. Each input in the sequence continuously activates the microcolumns to express itself through SPL. TPL activates cells in these minicolumns to express the context of current input and uses active cells to predict the follow-up content. TPL selects a small number of cells from active cells as learning cells to reduce the number of dendritic branches on cells. By adjusting the permanence of dendritic branches owned by these learning cells, TPL constructs the association between the input before and after and learns the sequence features. Active cell generation rule, learning cell generation

rule, and dendritic branch synaptic adjustment method are important factors affecting the performance of TPL. Next, from the two aspects of learning efficiency and learning effect, this paper analyzes the implementation details existing in TPL, such as the selection rules of learning cells and active cells and the updating rules of dendritic branches.

3.2.1. Learning Efficiency. In HTM, cells maintain a large number of dendritic branches, which record the relationship between cells. With these dendrite branches, HTM can save the sequence relationship between inputs so that the TPL can predict the future according to the current input. When a cell creates a new dendritic branch, its synapses are usually not connected. After TPL learns the input many times, the Hebbian rule makes the permanence of the dendrite branch increase or decrease. The cell can become the prediction cell of the associated active cell, and TPL completes the memory of the sequence only when the connected synapses in the dendrite branch reach the threshold. Therefore, the sequence needs to be learned repeatedly before being memorized by HTM.

Secondly, in the minicolumn activated at time t , TPL uses the following equation to generate active cells:

$$a_{ij}^t = \begin{cases} 1, & \text{if } j \in W^t \text{ and } \pi_{ij}^{t-1} = 1, \\ 1, & \text{if } j \in W^t \text{ and } \sum_i \pi_{ij}^{t-1} = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

If there are no predictive cells on an activated minicolumn at time t , TPL selects all cells on the minicolumn as active cells. TPL selects one of the cells as the learning cell to avoid creating too many dendritic branches in the learning process. TPL uses the following equation to select the learning cell on this minicolumn, which is called the best matching cell.

$$\begin{aligned} \dot{\text{cell}}_{ij}^t = 1 \text{ where } & j \in W^t \text{ and } \sum_i \pi_{ij}^{t-1} \\ & = 0 \text{ and } \exists k \left(\left\| \dot{D}_{ij}^k \circ A_1^{t-1} \right\| > \text{min Threshold} \right) \\ & \text{and } \max _ \text{match} \left(\left\| \dot{D}_{ij}^k \circ A_1^{t-1} \right\|_1 \right), \end{aligned} \quad (4)$$

where $\max_match(\|D_{ij}^k \circ A_1^{t-1}\|_1)$ determines whether cell_{ij} contains a dendritic branch that overlaps with the last active cells the most among all cells on j minicolumn. The min-Threshold is the minimum threshold of overlap values of dendrite branches, which is used to avoid some invalid dendritic branches from participating in the calculation.

TPL selects a dendrite branch on the learning cell and enhances its synaptic value to strengthen its association with active cells at time $t-1$. The selected dendrite branch D_{ij}^k should satisfy the following formula:

$$\|D_{ij}^k \circ A_1^{t-1}\|_1 = \max\left(\|D_{ij}^k \circ A_1^{t-1}\|_1\right). \quad (5)$$

In this case, with the unpredicted activated minicolumn, the adjusted dendrite branches have the following two situations:

- (1) The dendrite branch constructs the association between cell_i^t and Cell_i^{t-1} . TPL adjusts the synaptic permanence to enhance the association between the current input and the previous input.
- (2) The dendrite branch does not construct the association between cell_i^t and Cell_i^{t-1} . TPL adjusts the synaptic permanence to enhance the characteristics of a historical subsequence.

The use of equation (4) increases the probability of Case 2, which makes TPL give priority to learning the features of repeated subsequences. When TPL cannot find the learning cell with equation (4), TPL will randomly select one with the least utilization rate as the learning cell on the unpredicted activated minicolumn. TPL constructs a new dendrite branch on the cell and establishes the association with the learning cells at time $t-1$. In this way, TPL can learn and memorize the current new sequence feature and add it to the last learning content. Therefore, TPL needs to learn a sequence repeatedly before it can memorize and store all characteristics, even if TPL directly sets the initial synapse permanence of the newly created dendritic branch to the connected state.

3.2.2. Learning Effect. TPL gives priority to learning the features of repeated subsequences rather than those of the current input sequence. Therefore, after a round of training of the input sequence, TPL can only memorize and store the repeated segments. These memories are fragmentary. Compared with the complete sequence, there are many breakpoints in memory. Although more iterative training can repair these memory breakpoints, it also needs more time. By repairing memory breakpoints, fragment memories can string into a longer memory, which can be called the memory axis. However, some sequence fragments will be memorized again in the new context and form memory branches on the memory axis. The following experimental results can verify this conclusion.

Besides, there may be continuous same inputs in the sequence. Although the contents of these inputs are the same, their meanings are different due to their different positions. TPL selects different learning cells from the

activated minicolumns to represent the different contexts of input. The cells associated with learning cells also come from these minicolumns. When TPL learns the continuous same inputs, it is easy to make the input cycle predict itself. In this case, TPL will no longer learn the same new input, which will affect the training effect.

As shown in Figure 3, HTM has learned b for P times, and each b has different learning cells due to its different positions. When TPL learns the subsequent b , the active cells of the input may be a subset of the predicted cells, i.e., $A^t = A^t \circ \Pi^t$. This leads to the fact that the modified dendrite branches only strengthen the memory of history, and TPL no longer learns new repetitive content. Therefore, TPL cannot effectively learn repeated inputs, and HTM may only form one or more memory segments of the sequence.

When TPL learns a sequence, the memory generated by HTM usually has memory breakpoints and memory branches, which will affect the prediction effect. Memory breakpoints indicate that TPL cannot form a complete memory of a sequence, and they will reduce the recall of predictions. Memory branches will reduce the precision of prediction.

3.3. Cell Selection Rules Based on Location Awareness. In TPL_LA, the cell selection rules based on location awareness are designed. The rules reduce the range of active cells and no longer use the best matching cells. In this way, TPL_LA can distinguish the same input in different positions and give priority to learning the features of the current sequence.

When there are predictive cells on the activated minicolumns at time t , these cells contain active dendritic branches. Adjusting these dendrite branches will learn the sequence characteristics of the current position. TPL_LA uses the following equation to set the cell as a learning cell.

$$\text{cell}_{ij}^t = 1 \quad \text{where } j \in W^t \text{ and } \pi_{ij}^{t-1} = 1. \quad (6)$$

When there are not predictive cells on the activated minicolumns at time t , these cells do not contain active dendritic branches. TPL_LA uses the following equation to set the cell as a learning cell, where $\min_used(\text{cell}_{ij})$ determines whether cell_{ij} is the least used cell on j minicolumn.

$$\text{cell}_{ij}^t = 1 \quad \text{where } \min_used(\text{cell}_{ij}) \text{ and } (j \in W^t) \text{ and } \left(\sum_i \pi_{ij}^{t-1} = 0\right). \quad (7)$$

TPL_LA adds dendrite branches to the learning cell to construct the association with Cell_i^{t-1} to learn the sequence characteristics of the current position.

Finally, based on the selected learning cells, TPL_LA uses the following equation to create active cells.

$$A^t = \text{Cell}^t. \quad (8)$$

Different from the way TPL generates the learning cells and active cells, the algorithm no longer selects the best matching cells but randomly selects the cells with the least frequency as the learning cells with the unpredicted activated

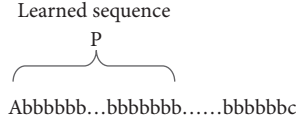


FIGURE 3: An example of repeated inputs in a sequence.

minicolumns. Therefore, TPL_LA can accurately express the context of the current input and learn the sequence features of the current position. Active cells usually express multiple contexts of input. The algorithm generates the active cells according to the learning cells, so it can accurately express the current context of the input. Such active cells reduce the prediction of historical content, which makes TPL learn the sequence characteristics of the current position and reduces the occurrence of memory breakpoints and memory branches. Besides, when TPL_LA learns the same continuous inputs, such active cells can also reduce the cycle prediction of the input and improve the learning effect of TPL_LA. The purpose of the proposed algorithm is to obtain the complete features of the rule. Every input owns different cell combinations to express the context through location awareness. To some extent, this increases the number of dendritic branches on the cell. This strategy will take up more storage space but improve the accuracy of prediction.

3.4. Dendrite Updating Rules Associated with Adjacent Inputs.

In TPL_LA, the updating rules of dendritic branches associated with adjacent inputs are designed, which makes TPL_LA quickly complete the learning of sequence features.

By adjusting the active dendrite branches of learning cells generated by equation (6), TPL_LA can learn the sequence features of the current position. TPL_LA uses the following equation to select the dendrite branches to be adjusted. The a_T is the activation threshold of dendritic branches.

$$\forall_{j \in w^t} (\pi_{ij}^{t-1} > 0) \text{ and } \tilde{D}_{ij}^k \circ A_1^{t-1} > a_T. \quad (9)$$

Reinforcing the above segments is to decrease all permanence values by a small value p^- and increase the permanence values corresponding to active presynaptic cells by a larger value p^+ [2].

$$D_{ij}^k = D_{ij}^k + p^+ (D_{ij}^k \circ A^{t-1}) - p^- D_{ij}^k. \quad (10)$$

By adding the active dendrite branches to the learning cells generated by equation (7), TPL_LA can establish the association between the current learning cell and $Cell^t$ and learn the sequence features of the current position. The algorithm adds a new dendrite branch D_{ij}^{new} according to the following equation. I_P is usually greater than the connectivity threshold of synapses.

$$D_{ij}^{\text{new}} = Cell^{t-1} \times I_P. \quad (11)$$

According to the above rules, either the selected active dendritic branches or the newly added dendritic branches construct the association between learning cells at the time

before and after. All of these make learning aim at the current sequence and reduce the interference of historical memory on current learning. When the synapse permanence is greater than or equal to the connectivity threshold, the synapse is connected. Therefore, using a bigger I_P , TPL_LA can quickly construct the temporal association of the sequence, extract the features of the memory sequence, and improve the learning efficiency.

3.5. *The Pseudocode of the Proposed Algorithm.* Based on Sections 3.3 and 3.4, the pseudocode of TPL_LA is shown as follows. (Algorithm 1).

TPL_LA modifies the rules of TPL. By sensing the position information of the input in the sequence, TPL_LA selects different learning cells to distinguish the same input. By reducing the range of active cells and no longer using the best matching cells as learning cells, TPL_LA only learns the sequence features at the current position. In this way, the memory of HTM rarely produces memory breakpoints and memory branches, and the algorithm can memorize all features of the sequence. Therefore, the learning effect of TPL_LA is highly improved. By increasing the synaptic permanence in the new dendrite branch of the learning cell, TPL_LA can learn the features of the sequences faster. Therefore, the algorithm not only guarantees the learning effect but also has high learning efficiency.

4. Experiments and Analysis

In the open-source framework of HTM, the prototype of TPL_LA is implemented. Different types of datasets are used to test and analyze it compared to TPL.

4.1. *Datasets and Evaluation Indicators.* In this paper, three datasets are used to test and analyze TPL_LA and TPL. Fifty digits after the π decimal point are used to construct the first dataset representing a random sequence. A sequence such as “211111..... 11113” is used as the second dataset to represent a sequence with continuously repeated inputs, whose length is set to 50. The third dataset is the statistical data of New York taxi passengers. From July 1, 2014, to January 31, 2015, the number of passengers per 30 minutes of New York taxi is used to construct the input sequence, with a total of 10320 data.

These three datasets are used to train the two algorithms, respectively, and the corresponding datasets are used for prediction tests. The test results are used to evaluate the learning effect of whether the two algorithms can obtain all features of the sequence. Three evaluation indicators, i.e., mean recall (MR), mean precision (MP), and F1, are used to evaluate the two algorithms. The recall^t describes whether the input is within the prediction range at time t . The precision^t describes the proportion of the input in the prediction range at time t . The evaluation indicators are calculated using the following equations:

```

(1) Get  $W^t$  of current input from spatial pool:
(2) Get  $A^{t-1} \Gamma^{t-1} Cell^{t-1}$  from last time
(3) For  $j = 0$  to  $N - 1$  do:
(4)   if  $j \in W^t$  then
(5)     for  $i = 0$  to  $M - 1$  do
(6)       if  $\pi_{ij}^{t-1} == 1$  then
(7)         set the current cell to  $cell_{i,j}^t$ 
(8)         if (learn) then
(9)           reinforce the segment of the cell;
(10)        endif
(11)       else
(12)         select the less used cell to  $cell_{i,j}^t$ 
(13)         if (learn)
(14)           create connected segment between  $cell_{i,j}^t$  and  $cell^{t-1}$ 
(15)         endif
(16)       endif
(17)     endfor
(18)   else
(19)     if  $\pi_{ij}^{t-1} == 1$  then
(20)       if (learn) punish the segment of the cell endif
(21)     endif
(22)   endif
(23) endfor
(24) if (learn)  $A^t = Cell^t$  endif
(25) Compute the  $\Gamma^t$ 

```

ALGORITHM 1: TPL_LA.

$$\text{recall}^t = \begin{cases} 1, & \text{if } A^t \circ \Gamma^{t-1} > a_T, \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{precision}^t = \begin{cases} \frac{\|A^t\|}{\|\Gamma^{t-1}\|}, & \text{if } \text{recall}^t = 1, \\ 0, & \text{otherwise,} \end{cases} \quad (12)$$

$$\text{MR} = \frac{1}{L} \sum \text{recall}^t,$$

$$\text{MP} = \frac{1}{L} \sum \text{precision}^t,$$

$$\text{F1} = \frac{1}{1/\text{MR} + 1/\text{MP}} = \frac{2 \times \text{MR} \times \text{MP}}{\text{MR} + \text{MP}}.$$

L is the length of a sequence. When the values of MR and MP are both large, it shows that the algorithm has a high degree of restoration for the input sequence. The larger the value of F1, the better the learning effect of the algorithm.

4.2. The Testing with Random Inputs. The sequence with random inputs is constructed using dataset 1. Figure 4 is a graphical representation of each input in dataset 1.

Dataset 1 has a small input range, so the prototype system uses Table 1 to set the relevant parameters.

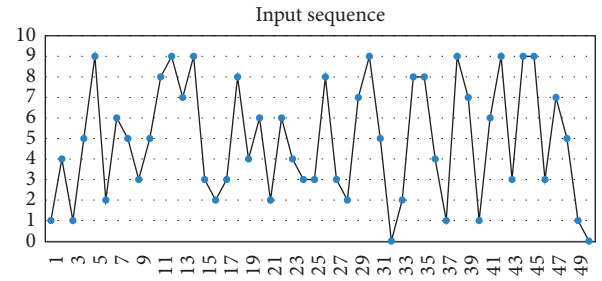


FIGURE 4: A sequence with random inputs.

TABLE 1: Parameters for a random sequence test.

Parameter	Parameter value	Meaning
C_D	128	Column dimensions
C_P_C	16	Cells per column
I_P	Pending	Initial permanence
C_P	0.8	Connected permanence
P_I	0.1	Permanence increment
P_D	0.05	Permanence decrement

4.2.1. The Comparison of Testing Results. When learning a sequence, TPL_LA needs larger initial value synapses. Therefore, I_P is set to the connectivity threshold of 0.8 and TPL_LA trains the sequence only once. TPL is tested in two cases. (1) TPL sets I_P to the connective threshold of 0.8 and trains the sequence once. (2) TPL sets I_P to the disconnected value of 0.7 and trains the sequence one and more times. Table 2 shows the testing results.

TABLE 2: Test results with random input.

Algorithm	I_P	Iterations	MR	MP	F1
TPL	0.7	1	0.32	0.25	0.28
TPL	0.7	10	0.90	0.74	0.81
TPL	0.7	50	0.96	0.85	0.90
TPL	0.7	100	0.98	0.89	0.93
TPL	0.8	1	0.82	0.59	0.69
TPL_LA	0.8	1	0.98	0.96	0.97

Through F1, it can be seen that TPL_LA is better than TPL. From the perspective of MR or MP, the results of the TPL_LA training sequence once are better than those of TPL in various situations. TPL_LA gives a distinct context to the input in different positions, so the recall and precision of the prediction are both the highest. In the learning process of TPL with connected I_P, repeated content is easy to form memory breakpoints. In addition, without repeated iterative learning, these breakpoints are difficult to repair, so the recall and precision of the prediction cannot reach the high point, which reduces the value of F1. In the learning process of TPL with unconnected I_P, nonrepetitive content is easy to form memory breakpoints. Repeated iterative learning can repair these memory breakpoints and produce some memory branches. Therefore, the learning result makes the recall of prediction very high but reduces the precision of prediction, which slightly reduces the value of F1. Although the TPL can also achieve a high approval of recall or precision through iterative learning, it needs more learning time. Therefore, TPL_LA has better learning efficiency and learning effect when learning a random sequence.

4.2.2. The Prediction State of Every Step. To better analyze the results in Table 2, the prediction states of each step when the two algorithms process dataset 1 are listed in Figure 5.

As can be seen from Figure 5(a), when TPL sets I_P disconnected and learns the sequence once, HTM only remembers the repeat patterns. Some patterns do not repeat in the dataset, which lets the synaptic value between the inputs of these patterns not reach the connectivity state, so HTM does not memorize these patterns. Compared with the complete sequence, there are many breakpoints in memory, so the recall and precision of prediction are very low. It can be seen from Figure 5(b) that when TPL sets I_P disconnected and learns the sequence 100 times, the memory in HTM can restore the complete sequence. Through iterative learning, TPL strengthens the disconnected synapses between inputs to the connected state, so these nonrepetitive patterns are also memorized. At the same time, the repeated patterns in different positions are given different contexts, so some memory branches are formed in HTM. Therefore, the recall of prediction is very high, and the precision of prediction is not high. It can be seen from Figure 5(c) that when TPL sets I_P connected and learns the sequence once, HTM memorizes long sequence fragments. However, memory breakpoints are formed at the position of each repetitive subsequence. The memory breakpoint will make the subsequent input have no context, which causes multiple

following predictions. As a result, the recall and precision of prediction are not very high. It can be seen from Figure 5(d) that when TPL_LA learns the sequence once, the memory in HTM can restore the complete sequence. The algorithm gives a distinct context to the input in different positions, so the memory in HTM has few memory breakpoints and memory branches. Therefore, the recall and precision of the prediction are both the highest.

4.3. The Testing with Repeated Inputs. The sequence with continuously repeated inputs is constructed using dataset 2. There are only three input types in the sequence, so the algorithms still use the parameters in Table 1.

4.3.1. The Comparison of Testing Results. For the sequence with continuously repeated inputs, the test is still carried out with the plan designed in Section 4.2.1. Table 3 shows the testing results.

Through F1, it can be seen that TPL_LA is better than TPL. From the perspective of MR or MP, the results of the TPL_LA training sequence once are better than those of TPL in various situations. TPL_LA also gives a distinct context to the input in different positions, so the recall and precision of the prediction are both the highest. In the learning process of TPL with connected I_P, continuous repeated content can be seen as the repeated iterative learning, which can repair the memory breakpoints. However, the continuous repeated content makes the active cell set the same as the predicted cell set, which prevents the algorithm from memorizing longer sequences. Therefore, the recall and precision of the prediction cannot reach a high point, which reduces the value of F1. In the learning process of TPL with unconnected I_P, repeated iterative learning can repair more memory breakpoints. However, because the active cell set is the same as the predicted cell set, the algorithm cannot memorize longer sequences. Therefore, the learning result makes the recall and precision of the prediction cannot reach the high point, which reduces the value of F1. Therefore, TPL_LA has better learning efficiency and learning effect when learning a sequence with continuously repeated inputs.

4.3.2. The Prediction State of Every Step. To better analyze the results in Table 3, the prediction states of each step when the two algorithms process dataset 2 are shown in Figure 5.

As can be seen from Figure 6(a), when TPL sets I_P disconnected and learns the sequence once, HTM only remembers a segment of the sequence. The existing algorithms will give priority to strengthening the repeated patterns. When the previous patterns are memorized, the algorithm will extend the memory content and increase the length of the memory axis. Since HTM only learns part of the sequence, the recall and precision of prediction are not high. As can be seen from Figure 6(b), when TPL sets I_P disconnected and learns the sequence 100 times, the memory in HTM still cannot restore the complete sequence. Through iterative learning, when the memorized sequence reaches a certain length, the active cells and prediction cells will be the

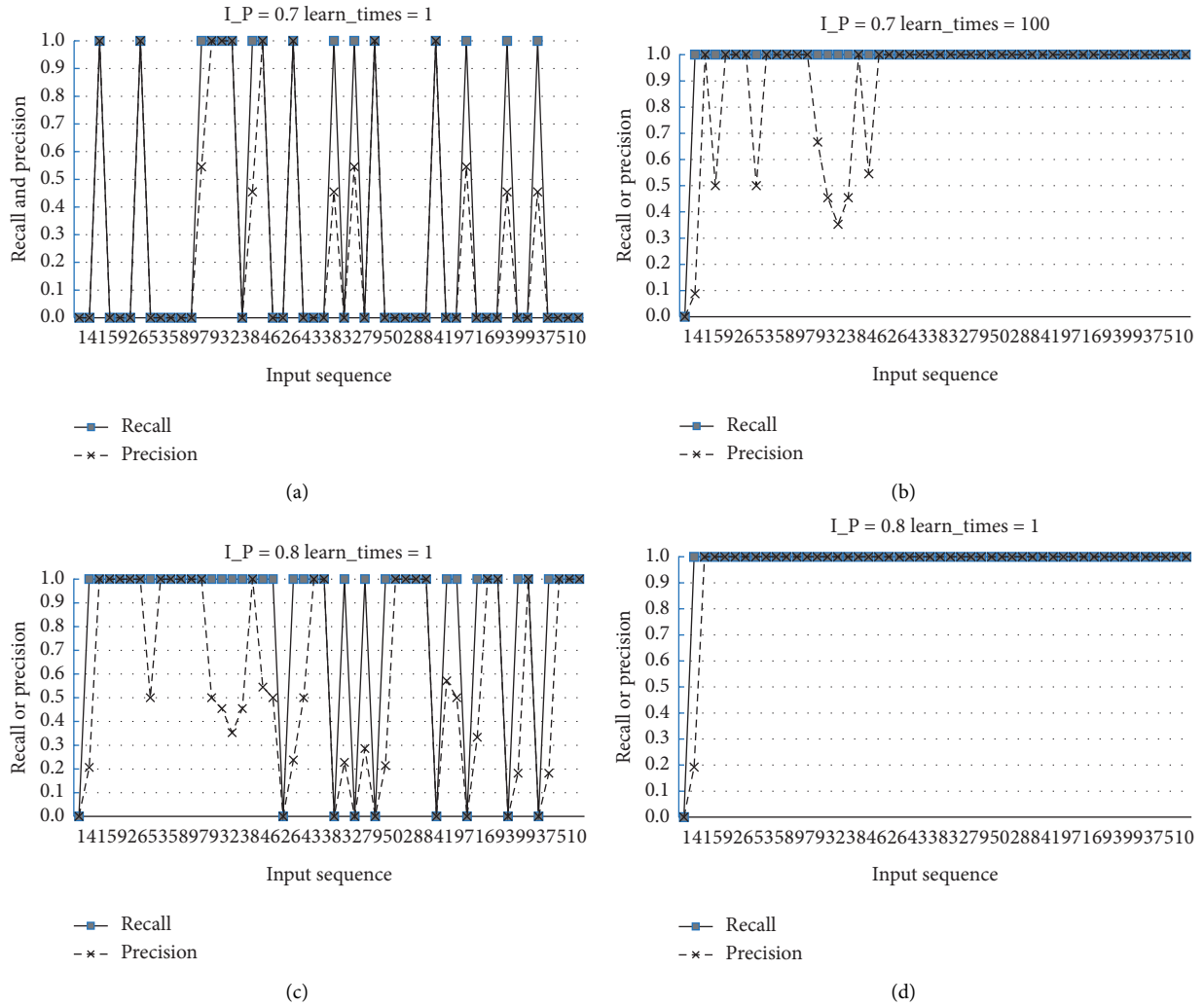


FIGURE 5: The prediction processes with a random sequence.

TABLE 3: Test results with repeated inputs.

Algorithm	I_P	Iterations	MR	MP	F1
TPL	0.7	1	0.82	0.82	0.82
TPL	0.7	10	0.88	0.82	0.85
TPL	0.7	20	0.90	0.80	0.85
TPL	0.7	50	0.90	0.80	0.85
TPL	0.8	1	0.88	0.86	0.87
TPL_LA	0.8	1	0.98	0.98	0.98

same, which leads to the algorithm not learning new repeated content. At the same time, the repeated pattern, especially the “13,” owns different contexts, so some memory branches are formed in HTM. Therefore, the recall of prediction is not improved much, and the precision of prediction is decreased. As can be seen from Figure 6(c), when TPL sets I_P connected and learns the sequence once, HTM still only remembers a segment of the sequence. The reason is also that when the repeat sequence is memorized to a certain length, the active cells and the prediction cells will be the same, resulting in the fact that the algorithm could not

learn new content. The algorithm learns the sequence “13” only once, so the pattern has only one context. Compared with the results of Figure 6(c), the recall and the precision of the prediction are not much different. As can be seen from Figure 6(d), when TPL_LA learns the sequence once, the memory in HTM can restore the complete sequence. The algorithm gives different contexts to the input in distinct positions, so the memory in HTM has few memory breakpoints and memory branches. Therefore, the recall and precision of the prediction are both the highest.

4.4. The Testing of the Taxi Passenger Flow of New York.

The real data sequence is constructed by dataset 3. Because the dataset is much larger than the first two datasets, TPL_LA and TPL use Table 4 to set the parameters.

In this testing, TPL_LA set I_P to 0.8. TPL is tested in two cases. (1) TPL sets the I_P to 0.8. (2) TPL sets I_P to 0.7. Each algorithm only learns the data set once under the corresponding configuration.

The dataset has 10320 data and 8090 different numbers. This dataset may be seen as the first level dataset, which

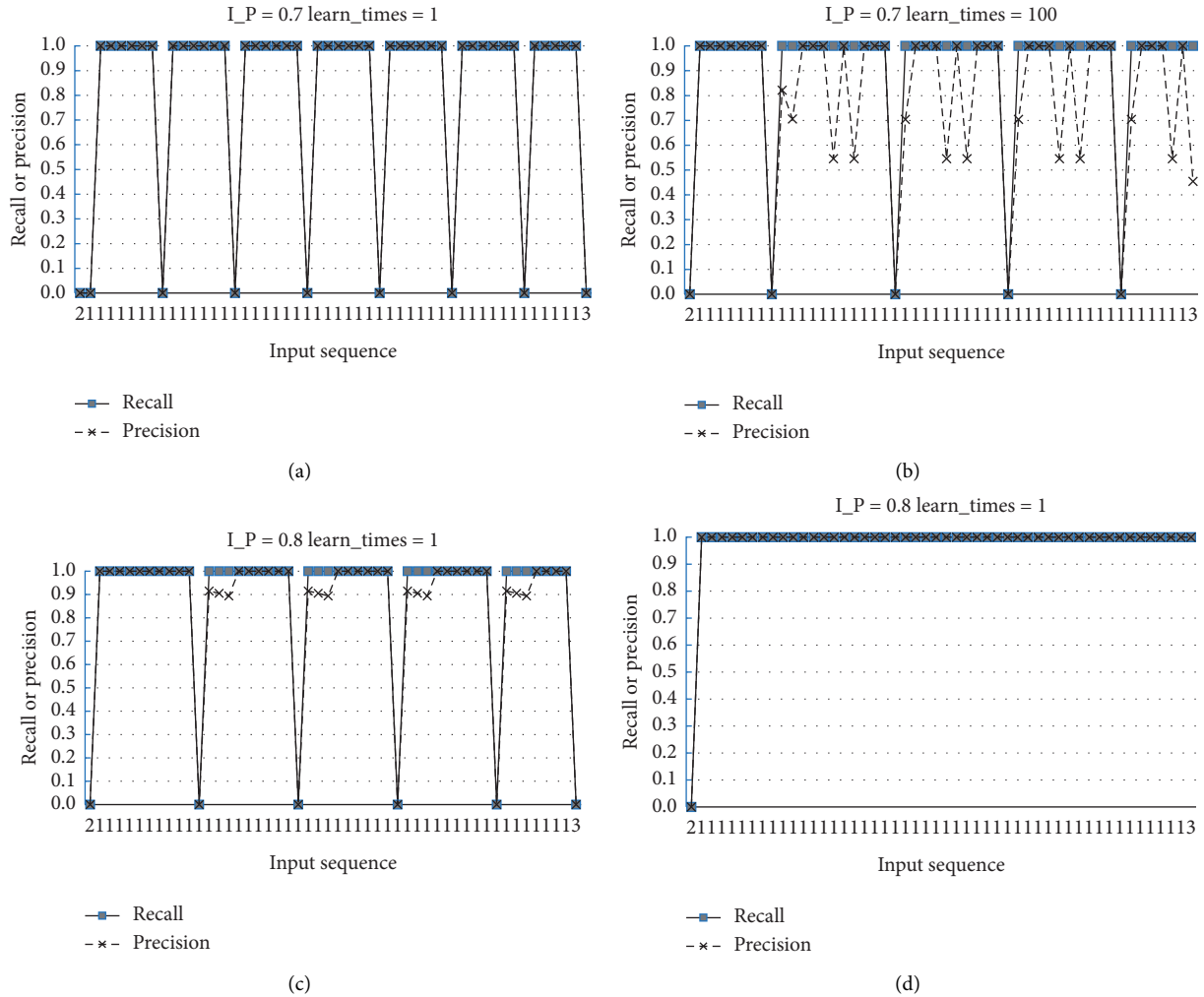


FIGURE 6: Processes of the continuous repeated sequence.

TABLE 4: Parameters for taxi passenger flow testing.

Parameter	Parameter value	Meaning
C_D	2048	Column dimensions
C_P_C	16	Cells per column
I_P	Pending	Initial permanence
C_P	0.8	Connected permanence
P_I	0.1	Permanence increment
P_D	0.05	Permanence decrement

includes less repeated subsequence. Table 5 shows the testing results on the first level dataset.

To further study the performance of the algorithm, each data in the original dataset is divided by 10. In this way, the number of repeats in the sequence can be increased by increasing the unit of measurement of the data. This dataset may be seen as the second level dataset, which has 10320 data and 2470 different numbers. Table 6 shows the testing results on the second level dataset.

Go further, each data in the original dataset is divided by 100. This dataset may be seen as the third level dataset, which

has 10320 data and 287 different numbers. The data set contains more repetitive subsequences. Table 7 shows the testing results on the third level dataset.

From the indicators in Tables 5–7, it can be seen that TPL_LA always has the highest MR and MP. TPL_LA gives different contexts to the input in different positions and can enable HTM to remember all features of the sequence by learning them once. The memory in HTM can completely restore the content of the sequence, no matter how many repetitive patterns are included in the sequence. TPL_LA has good learning efficiency and learning effect. If only learning the sequence once, TPL with connected I_P can get a good learning effect on the sequence without repeated content, but the learning effect will decline when there are repeated contents in the sequence. In this setting, repeated content is easy to form memory breakpoints. On the other hand, if only learning the sequence once, TPL with disconnected I_P can hardly remember the characteristics of the sequence without repeated contents. In this setting, the synapses between data cannot reach the connected state, and the temporal relationship between data cannot be formed in HTM. Only

TABLE 5: The test results of the taxi passenger flow (first level dataset).

Algorithm	INITIAL_PERMANENCE	Iterations	MR	MP	F1
TPL	0.7	1	0.0	0.0	0.0
TPL	0.8	1	0.95697	0.95697	0.95697
TPL_LA	0.8	1	0.9999	0.9999	0.9999

TABLE 6: The test results of the taxi passenger flow (second level dataset).

Algorithm	INITIAL_PERMANENCE	Iterations	MR	MP	F1
TPL	0.7	1	0.01337	0.01166	0.01246
TPL	0.8	1	0.67955	0.67142	0.6755
TPL_LA	0.8	1	0.9999	0.9999	0.9999

TABLE 7: The test results of the taxi passenger flow (third level dataset).

Algorithm	INITIAL_PERMANENCE	Iterations	MR	MP	F1
TPL	0.7	1	0.35193	0.0802	0.13063
TPL	0.8	1	0.4718	0.19436	0.2753
TPL_LA	0.8	1	0.9999	0.9999	0.9999

when there are repeated contents in the sequence or HTM learns the sequence more times can the algorithm extract and memorize the features of the sequence.

5. Conclusions

Through the analysis of TPL, the authors explain the reasons for the low learning efficiency and poor learning effect of TPL and propose an efficient temporal pool learning algorithm based on location awareness. After one training of TPL_LA, the sequence can be effectively memorized by HTM, which shows that the algorithm has good learning efficiency and learning effect.

In the following research, the authors intend to add more input feature information to the TPL_LA and modify the learning mechanism of the proposed algorithm, which could enhance the smoothness of the memory, make the predictive generalization higher, and improve the utilization rate of the network space of the HTM. In this way, the proposed algorithm can be better used in the intelligent network in the future.

Data Availability

The artificial data used to support the findings of this study are included within the article. The time-period statistics of New York taxi passengers can be found at <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research work was supported in part by the National Natural Science Foundation of China (grant no. 61806086), in part by the Project of National Key R&D Program of

China (2018YFB0804204), and in part by Jiangsu Provincial doctoral Innovation Fund (CXZZ13_0689).

References

- [1] J. Hawkins, S. Ahmad, and D. Dubinsky, Hierarchical temporal memory including htm cortical learning algorithms," http://numenta.org/resources/HTM_CorticalLearningAlgorithms.pdf, 2011.
- [2] J. Hawkins and S. Ahmad, "Why neurons have thousands of synapses, a theory of sequence memory in neocortex," *Frontiers in Neural Circuits*, vol. 10, pp. 1–13, 2016.
- [3] D. George, "How the brain might work: a hierarchical and temporal model for learning and recognition," Ph.D Thesis, Stanford University, Kunnamangalam, Kerala, 2008.
- [4] J. Hawkins, S. Ahmad, S. Purdy, and A. Lavin, "Biological and machine intelligence (Bami)," <http://numenta.com/biological-and-machine-intelligence/>, 2016.
- [5] S. Ahmad and J. Hawkins, "Properties of sparse distributed representations and their application to hierarchical temporal memory," 2015, <https://arxiv.org/abs/1503.07469>.
- [6] S. Lattner, "Hierarchical temporal memory-investigations, ideas, and experiments," M.Sc. thesis, Johannes Kepler Universität, Linz, Austria, 2014.
- [7] F. Byrne, "Encoding reality: prediction-assisted cortical learning algorithm in hierarchical temporal memory," 2015, <https://arxiv.org/abs/1509.08255>.
- [8] M. Leake, L. Xia, K. Rocki, and W. Imano, "A probabilistic view of the spatial pooler in hierarchical temporal memory," *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, vol. 9, pp. 1111–1118, 2015.
- [9] J. Mnatzaganian, E. Fokoué, and D. Kudithipudi, "A mathematical formalization of hierarchical temporal memory's spatial pooler," *Frontiers in Robotics and AI*, vol. 3, p. 81, 2017.
- [10] L. Li, T. Zou, T. Cai, D. Niu, and Y. Zhu, "A fast spatial pool learning algorithm of hierarchical temporal memory based on minicolumn's self-nomination," *Computational Intelligence*

- and Neuroscience*, vol. 2021, Article ID 6680833, 13 pages, 2021.
- [11] D. Fan, M. Sharad, A. Sengupta, and K. Roy, "Hierarchical temporal memory based on spin-neurons and resistive memory for energy-efficient brain-inspired computing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 9, pp. 1907–1919, 2016.
 - [12] O. Krestinskaya, I. Dolzhikova, and A. P. James, "Hierarchical temporal memory using memristor networks: a survey," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 5, pp. 380–395, 2018.
 - [13] O. Krestinskaya, T. Ibrayev, and A. P. James, "Hierarchical temporal memory features with memristor Logic circuits for pattern recognition," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 6, pp. 1143–1156, 2018.
 - [14] D. Maltoni, "Pattern recognition by hierarchical temporal memory," *SSRN Electronic Journal*, 2011.
 - [15] A. Irmanova, T. Ibrayev, and A. P. James, "Discrete-level memristive circuits for HTM-based spatiotemporal data classification system," *IET Cyber-Physical Systems: Theory & Applications*, vol. 3, no. 1, pp. 34–43, 2018.
 - [16] T. Ibrayev, U. Myrzakhan, O. Krestinskaya, A. Irmanova, and A. P. James, "On-chip face recognition system design with memristive Hierarchical Temporal Memory," *Journal of Intelligent & Fuzzy Systems*, vol. 34, no. 3, pp. 1393–1402, 2018.
 - [17] M. Wielgosz and M. Pietroń, "Using spatial pooler of hierarchical temporal memory to classify noisy videos with predefined complexity," *Neurocomputing*, vol. 240, pp. 84–97, 2017.
 - [18] A. P. James, I. Fedorova, T. Ibrayev, and D. Kudithipudi, "HTM spatial pooler with memristor crossbar circuits for sparse biometric recognition," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 11, no. 3, pp. 640–651, 2017.
 - [19] R. Skoviera, I. Bajla, and J. Skovierová, "Object recognition in clutter color images using Hierarchical Temporal Memory combined with salient-region detection," *Neurocomputing*, vol. 307, pp. 172–183, 2018.
 - [20] F. Asgari and A. Salehi, "The biologically inspired hierarchical temporal memory model for farsi handwritten digit and Letter recognition," *International Journal of Computer Applications*, vol. 129, no. 16, pp. 6–11, 2015.
 - [21] Y. Zou, Y. Shi, Y. Wang, Y. Shu, Q. Yuan, and Y. Tian, "Hierarchical temporal memory enhanced one-shot distance learning for action recognition," in *Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, San Diego, CA, USA, July 2018.
 - [22] J. Luo and T. Tjahjadi, "Gait recognition and understanding based on hierarchical temporal memory using 3D gait semantic folding," *Sensors (Basel, Switzerland)*, vol. 20, 2020.
 - [23] A. Irmanova, "Image based HTM word recognizer for language processing," in *Proceedings of the 2018 IEEE International Conference on Consumer Electronics—Asia (ICCE-Asia)*, pp. 206–212, JeJu, South Korea, June 2018.
 - [24] Y. Bouhadjar, M. Diesmann, R. Waser, D. J. Wouters, and T. Tetzlaff, "Constraints on sequence processing speed in biological neuronal networks," in *Proceedings of the International Conference on Neuromorphic Systems (ICONS '19)*, Knoxville, TN, USA, 2019.
 - [25] N. O. El-Ganainy, I. Balasingham, P. S. Halvorsen, and L. Arne Rosseland, "On the performance of hierarchical temporal memory predictions of medical streams in real time," in *Proceedings of the 2019 13th International Symposium on Medical Information and Communication Technology (ISMICT)*, pp. 1–6, Oslo, Norway, May 2019.
 - [26] S. Suzugamine, T. Aoki, K. Takadama, and H. Sato, "Self-structured cortical learning algorithm by dynamically adjusting columns and cells," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 24, no. 2, pp. 185–198, 2020.
 - [27] A. M. Ziyarah and D. Kudithipudi, "Neuromorphic architecture for the hierarchical temporal memory," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 1, pp. 4–14, 2019.
 - [28] A. Anandharaj and P. B. Sivakumar, "Anomaly detection in time series data using hierarchical temporal memory model," in *Proceedings of the 2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1287–1292, Coimbatore, India, June 2019.
 - [29] T. Charrad, K. Noura, and A. Ferchichi, "Deconstructing RPCs," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6S2, pp. 312–315, 2019.
 - [30] H. Luo, S. Jia, and W. Zhang, "Hierarchical temporal memory based anomaly detection for hydrological monitoring of unmanned surface vehicle," in *Proceedings of the 2019 IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP)*, pp. 420–424, Weihai, China, September 2019.
 - [31] W. Midani, Z. Fki, and M. BenAyed, "Online anomaly detection in ECG signal using hierarchical temporal memory," in *Proceedings of the 2019 fifth international Conference on Advances in Biomedical engineering (ICABME)*, pp. 1–4, Weihai, China, September 2019.
 - [32] A. Bamaqa, M. Sedky, T. Bosakowski, and B. B. Bastaki, "Anomaly detection using hierarchical temporal memory (HTM) in crowd management," in *Proceedings of the 2020 4th International Conference on Cloud and Big Data Computing*, London, UK, August 2020.

Research Article

Segmentation Technology of Nucleus Image Based on U-Net Network

Jie Fang ¹, QingBiao Zhou ¹ and Shuxia Wang ²

¹Zhejiang Industry Polytechnic College, Shaoxing 312000, Zhejiang, China

²School of Computing, Tonghua Normal University, Tonghua 134001, Jilin, China

Correspondence should be addressed to QingBiao Zhou; qingbiaozhou@163.com

Received 19 April 2021; Revised 28 May 2021; Accepted 2 June 2021; Published 11 June 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Jie Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To solve the problems of rough edge and poor segmentation accuracy of traditional neural networks in small nucleus image segmentation, a nucleus image segmentation technology based on U-Net network is proposed. First, the U-Net network is used to segment the nucleus image, which stitches the feature images in the channel dimension to achieve feature fusion, and the skip structure is used to combine the low- and high-level features. Then, the subregional average pooling is proposed to improve the global average pooling in the attention module, and an attention channel expansion module is designed to improve the accuracy of image segmentation. Finally, the improved attention module is integrated into the U-Net network to achieve accurate segmentation of the nuclear image. Based on the Python platform, the experimental results show that the proposed segmentation technology can achieve fast convergence, and the mean intersection over union (MIoU) is 85.02%, which is better than other comparison technologies and has a good application prospect.

1. Introduction

With the development of medicine, more and more medical imaging images need to be processed, and image processing technology has become more and more important [1]. Traditional medical imaging image processing and analysis only rely on the doctor's experience, which not only wastes manpower but also affects the accuracy rate because the doctor's experience and physical condition affect judgment result. Therefore, the breakthrough of automated medical image processing technology has a very critical role in improving the efficiency of medical diagnosis [2]. Medical image segmentation is an important task. Moreover, many other related tasks in medical image processing require image segmentation in advance. Medical image segmentation generally refers to the extraction of certain target regions in the entire image in some way, such as cell nuclei, an organ, or tissue [3, 4]. The results of medical image segmentation usually have no intersection in each area. Moreover, each segmented area has a certain similarity in its interior [5]. Compared with image segmentation of natural

scenes with clear outlines, medical images have great particularities. The complexity of the medical image itself causes the separation between its components to be blurred, and the boundaries between the components are not clear enough [6].

The method of medical image segmentation has developed from the initial manual segmentation to semi-automatic segmentation and then to the most recent fully automatic segmentation [7, 8]. With the gradual in-depth study of medical image segmentation by domestic and foreign researchers, many research results based on traditional algorithms have been obtained, which can be divided into three categories [9]. One is an algorithm that uses the discontinuity of boundary information to perform segmentation, such as surface fitting, parallel differential operator, deformation model, and so on [10]. The second is an algorithm for segmentation using the similarity of different regions of the image, such as region growing algorithm, threshold method, classification and clustering algorithm, and statistical-related algorithm [11]. The third is an algorithm that combines the discontinuity of boundary

information and the similarity of different regions of the image [12]. Reference [13] proposed an immune system programming (ISP) image segmentation algorithm based on a new evolutionary algorithm combined with region growth technology. The ISP algorithm with a tree data structure can segment medical images better. However, the actual area growth technology does not consider the complexity of the image boundary, and the accuracy rate needs to be optimized. Reference [14] proposed an improved multi-level threshold image segmentation method based on differential evolution. The efficiency of different parts of the allocation differential evolution algorithm is evaluated by measuring the quality of the candidate solutions, so as to generate the optimal solution of the allocation population to improve the efficiency of the algorithm when the number of thresholds increases. Reference [15] proposed a non-revisited quantum behavior particle swarm algorithm. Among them, the use of a refined search method overcomes the shortcomings of the original search method, reduces the computational cost, and has better effectiveness and robustness. Reference [16] proposed an active contour segmentation method for morphological medical images with an automatic initialization function. This method has a low computational cost, good robustness, and a high degree of automation. However, the settings before initialization are more complicated, and the adaptability to different application scenarios is low.

Traditional algorithms have obvious limitations. In recent years, deep learning methods have developed rapidly and are widely used, such as recurrent neural networks, restricted Boltzmann machines, and convolutional neural networks [17]. Reference [18] proposed a deep belief network brain tumor image segmentation method based on harmonious cuckoo search. By integrating Bayesian fuzzy clustering and the active contour model, better accuracy is obtained, but the computational efficiency is not high. Reference [19] achieved high-precision tumor segmentation in the segmentation process of the fuzzy mean clustering algorithm by extracting the features of the gray co-occurrence matrix and the grayscale run-length matrix. However, the steps in the feature extraction stage are more complicated, and the segmentation efficiency needs to be improved. Reference [20] developed a multi-graph-based tag fusion high-order feature learning framework. Fusion of the mean-covariance limited Boltzmann machine and high-level image features to segment structural brain images. Reference [21] used the evolutionary computing power potential of dense blocks and residual blocks to propose an automatic evolution model for medical image segmentation. Good results have been achieved, but there remain some difficulties for image segmentation of complex nuclei.

Aiming at the fact that the existing segmentation methods are difficult to apply to the image segmentation of the nucleus in the medical field, a U-Net-based cell nucleus image segmentation technology is proposed. Compared with traditional medical image segmentation methods, its innovations are

- (1) To solve the problems of poor segmentation of small nucleus, rough edges, and under- and over-segmentation, the U-Net network is used for image segmentation. It stitches feature maps in the channel dimension to achieve feature fusion and uses a skip structure to combine low- and high-level features to ensure the segmentation effect of the nucleus.
- (2) Due to the commonly used global mean pooling method, the extracted channel attention information is weak in interpretability, and the information is rough. Therefore, the subregional average pooling method is used in the attention module instead of the global average pooling method.

The structure of this paper is as follows: Section 1 introduces the significance and research status of medical image segmentation and summarizes the innovation points of the proposed segmentation network. Section 2 introduces the U-Net network in detail, as well as the attention mechanism and its improvement methods, thereby designing a complete nucleus segmentation network structure. Section 3 conducts experiments and evaluates the results to demonstrate that the proposed segmentation network has good feasibility and effectiveness. Finally, the full text is summarized and prospected.

2. Theory and Method

2.1. U-Net. Because the semantic segmentation effect of fully convolutional networks (FCN) is relatively rough, and the U-Net network, as a further extension of the FCN, has become the cornerstone of medical image segmentation [22, 23]. U-Net is a semantic segmentation network proposed by Olaf Ronneberger in 2015. In the upsampling process, the form of downsampling is matched to keep it consistent [24]. On this basis, a large number of feature maps in the downsampling stage are added to the upsampling to fill in the information lost in the calculation process. Its structure is shown in Figure 1.

U-Net includes a contraction path and an expansion path. The left side of Figure 1 is a contraction path, that is, downsampling, including two 3×3 convolutional layers and a 2×2 maximum pooling layer with a stride of 2. The activation function is rectified linear unit (ReLU) [25]. The classic image classification network with the fully connected layer removed is usually used. It performs a convolution kernel pooling operation on the original input picture, which can obtain contextual semantic information to solve the classification problem in image segmentation. On the right is an expansion path, that is, upsampling, which can locate segmentation tasks [26]. Up- and downsampling are symmetrical. First, a 2×2 convolutional layer is connected to reduce the number of feature channels, and then two 3×3 convolutions are used. Finally, there is a convolutional layer of 1×1 to map the required number of feature vectors, and the convolution is an unpadding structure.

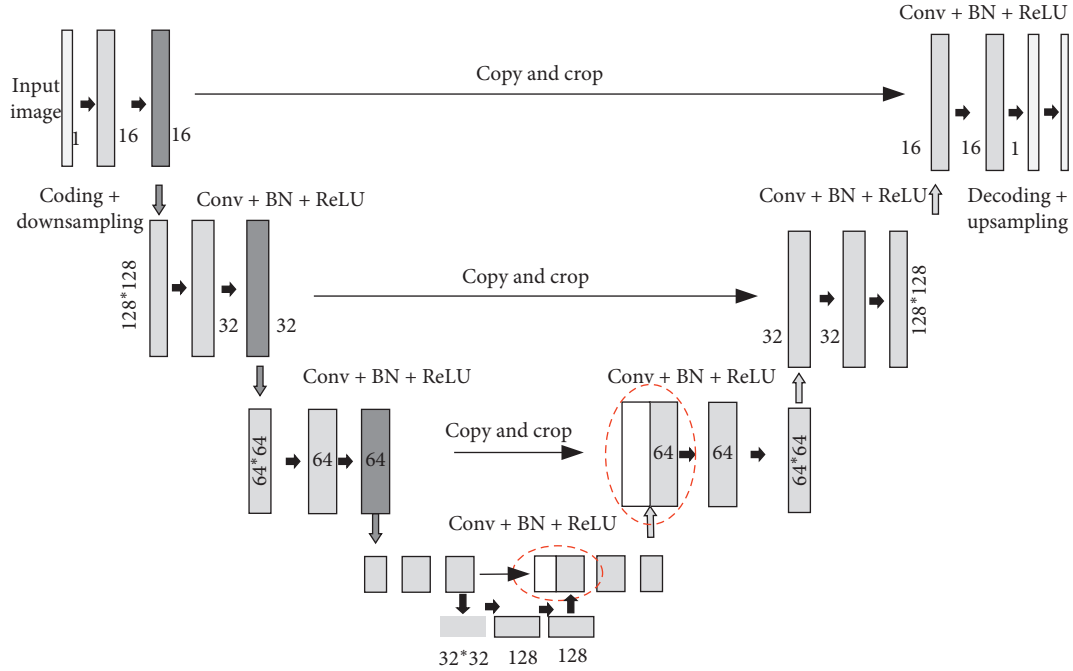


FIGURE 1: U-Net network structure.

The U-Net network has a major change in structure, and it builds more characteristic channels through skip connections during upsampling. Moreover, U-Net and FCN use different methods for feature fusion. FCN adds the feature maps point by point, while U-Net splices the feature maps to make them have more channels. The model training requires fewer data sets, can converge on a small amount of data, and can quickly obtain results when performing image segmentation.

The energy function of network training uses weighted cross-entropy, which is calculated as follows:

$$E = \sum_{x \in \phi} w(x) \log(p_k(x)),$$

$$p_k(x) = \frac{\exp(\sigma_k(x))}{\sum_{k'}^K \exp(\sigma_{k'}(x))},$$
(1)

where $\sigma_k(x)$ represents the activation function of the k -th feature channel at the x -th pixel. k is the number of categories. $p_k(x)$ is the approximate maximum function. ϕ is the image set. $w(x)$ is the importance of pixels in the training composition; the more the importance, the greater the weight.

U-Net has special advantages in processing medical images. To solve the problem of the lack of sample images in medical images, elastic deformation is used to complete data enhancement. Elastic deformation is a relatively common type of deformation in actual cells, so it is very suitable for medical image processing [27, 28]. The algorithm of data enhancement is adopted to make the neural network model learn the invariance of elastic deformation so that the network can have good elastic deformation adaptability when the data set is small. It can correctly complete the

segmentation when encountering the elastically deformed medical image.

2.2. Attention Module. To improve the segmentation speed of the network, a lightweight network is used as a feature extraction network, but there is a certain loss in the accuracy. In order to improve the accuracy, adding an attention module to the U-Net helps enhance the feature expression of the model [29, 30]. This module integrates different information and improves the understanding of the model, which is similar to the attention mechanism of human vision. There are two types of human visual attention mechanisms: Bottom-up data-driven attention mechanism and top-down task-driven target attention mechanism [31]. Both mechanisms can learn the parts required by the task from a large amount of data. The proposed network uses a bottom-up data-driven attention mechanism [32]. The attention module starts from the relationship between feature channels and considers the interdependence factors between feature channels. Through the self-learning of the network, the features that have little effect on the current segmentation are effectively suppressed, and the weight of beneficial features is enhanced. Its module structure is shown in Figure 2.

The attention module first performs global average pooling (GAP) on the feature map of each channel to obtain the vector of $1 \times 1 \times K$, and then performs two fully connected (FC) layer conversions. To suppress the complexity of the model, dimensionality reduction and increase are performed between the two FC layer conversions, which is similar to a “bottleneck”, and Sigmoid and ReLU activation functions are used.

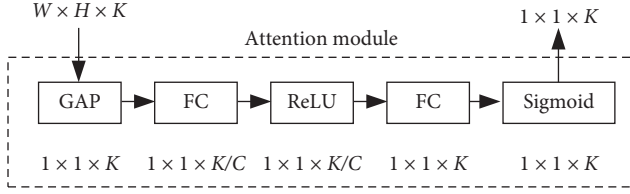


FIGURE 2: The structure of attention module.

Image semantic segmentation tasks are often aimed at the processing of complex scenes. There are often multiple objects in the image; the object types and sizes are different; and the spatial scene distribution is complex. Using GAP directly means simply treating each channel as a single-category problem, and GAP has no parameters to learn from [33]. The channel information obtained by this method may be too rough to explain the meaning of the channel better [34]. To solve the problem and learn the spatial distribution of the image better, a GAP improvement method is proposed, namely, subregional average pooling (SAP). GAP directly transforms a channel feature with a dimension of $H \times W \times 1$ into a $1 \times 1 \times 1$ size feature, and the process of SAP processing is shown in Figure 3.

SAP first transforms a channel feature with a shape of $H \times W \times 1$ into a new feature map with a dimensional shape of $m \times n \times 1$ after an adaptive mean pooling operation so that the obtained channel feature retains some spatial information. Then, a convolution operation with a kernel size of $m \times n$ is used to convert the obtained new channel feature map into a channel feature with a dimension of $1 \times 1 \times 1$. The convolution operation can learn the characteristics of the image after pooling and get its spatial characteristics. It should be noted that in the experiment, in order to facilitate the calculation and not increase the amount of calculation too much, the parameter is set to $m = n = 4$, and it is better not to be too large for m and n , generally less than or equal to 7.

In order to allow low-level channel information to be transmitted to higher-level channels, low- and high-level channel information are aligned at the same time. An attention channel expansion module CE is designed in the proposed network, as shown in Figure 4.

Among them, a 1×1 convolution operation is used to increase the number of channels, and then the Sigmoid function is used [35]. The high-level channel attention weight and the low-level expanded channel attention weight are added to obtain the updated channel attention weight. Finally, the recalibrated channel weight is multiplied by the high-level features. It should be noted that the range of the updated channel weight is $(0, 2)$ instead of the usual $(0, 1)$. This will allow the weight to not only reduce the original feature value but also expand the feature value [36].

In order to facilitate the description, the input feature map is denoted as $U = [u_1, u_2, \dots, u_C]$, $u_i \in \mathcal{R}^{H \times W}$ represents the i -th channel map in U . $f \in \mathcal{R}^{m \times n \times C}$ represents the feature map of U after the pooling operation, and its dimension is $m \times n \times C$. $u_c^g \in \mathcal{R}^{h \times w}$ indicates that the pooling operation corresponds to the characteristic subregion g of

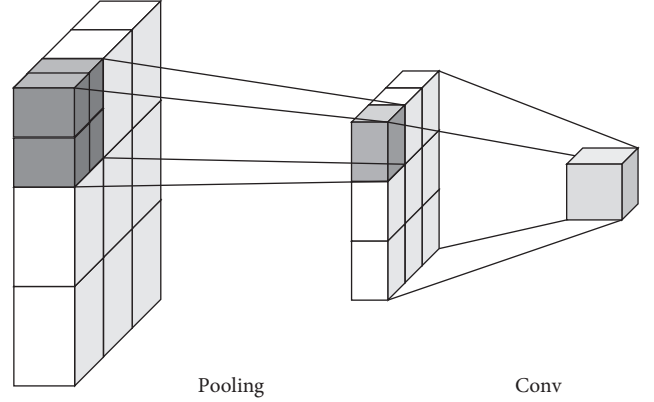


FIGURE 3: SAP process.

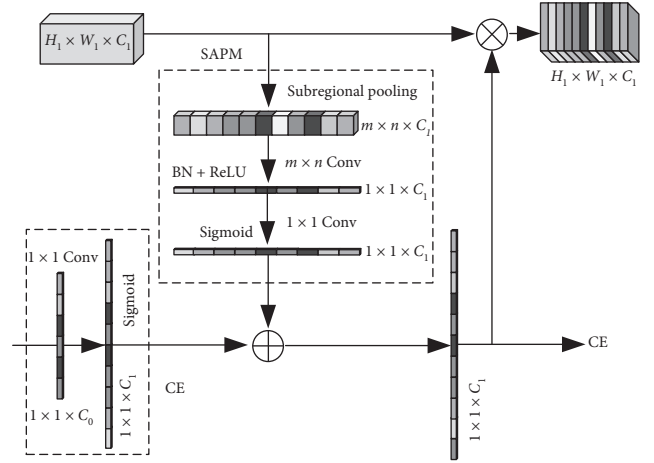


FIGURE 4: Improved channel attention model.

the c -th channel graph in U . The size of the subregion is $h \times w$. Then the expression of the pooling operation can be obtained as follows:

$$f_c^{(m,n)} = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w u_c^g(i, j). \quad (2)$$

Next, it is a convolution calculation, defining symbol $*$ represents the convolution, $d_c \in \mathcal{R}^{m \times n}$ represents the conjunction core corresponding to the f_c , and the result F_c is represented as follows:

$$F_c = d_c * f_c. \quad (3)$$

The activation function ReLU is used after batch normalization (BN) for $F = [F_1, F_2, \dots, F_C]$. Here, φ represents the BN operation, and σ represents the ReLU function.

$$\tilde{F} = \sigma(\varphi(F)). \quad (4)$$

Then, the obtained channel feature \tilde{F} is then connected to the convolutional calculation s_r of 1×1 , and the function is activated using the Sigmoid, where $r \in [1, C]$. Remember the Sigmoid function is δ , the output δ_{SAP} of the SAP module is as follows:

$$\delta_{\text{sapm}} = [\delta_1, \delta_2, \dots, \delta_r, \dots, \delta_C],$$

$$\delta_r = \delta \left(\sum_{i=1}^C s_r * \tilde{F}_i \right). \quad (5)$$

Finally, the high-order channel attention distribution obtained by defining SAP module is $\delta'(\cdot)$. The low-order channel attention distribution of the CE module is $\delta''(\cdot)$. Then we can get the channel attention distribution δ_F for updating higher-order features.

$$\delta_F = \delta''(\cdot) + \delta'(\cdot). \quad (6)$$

After the above steps, a recalibrated feature \hat{U} is $\hat{U} = U\delta_F$.

2.3. Network Structure Design. In the U-NET network, the improved attention module is integrated for nucleus image segmentation. Its network structure is shown in Figure 5.

Three layers of 3×3 convolution are used in the network; the last classification layer is removed; and the step size of the maximum pooling layer is changed from 2 to 1. At the same time, the ordinary 3×3 convolution is changed to the expansion convolution with the expansion rate of 2 so that the resolution of the output feature map is equal to 1/16 of the size of the input image. An improved atrous spatial pyramid pooling (ASPP) module is added at the top of the network, and its output channel number is 512. After 4 times of upsampling, the output of the features by ASPP is added to a low-level feature that has undergone a 3×3 convolution and has the same dimension. Then connect a 3×3 convolution operation for feature fusion. Finally, it is upsampled to restore the original image size. Among them, the SAP module is used to obtain the initial channel weight information, and the CE module is used to expand the low-level channel attention.

Finally, each layer of the network is connected by skip connections, where connections exist in all layers. And in the pooling method, ASPP is used. ASPP provides a multi-scale information model, which adds a dilated convolution with different expansion rates on the basis of spatial pyramid pooling to capture a wide range of contexts. SAP is used to combine image features to increase the global context.

3. Experimental Results and Analysis

The network is built through the TensorFlow deep learning framework released by Google. The GPU model used is RTX 2080Ti, and the card memory size is 11 GB. The main information of the experiment is shown in Table 1.

3.1. Network Parameter Setting. In the experiment, ASPP was changed to 4 parallel 3×3 expansion convolution operations, rates=(1, 6, 12, 18). At the same time, during network training, random gradient descent (SGD) is used for parameter optimization. Its momentum parameter is set to 0.98, and the weight decay rate is 1×10^{-4} . The learning rate is a reduction strategy of the initial learning rate $\gamma_0 =$

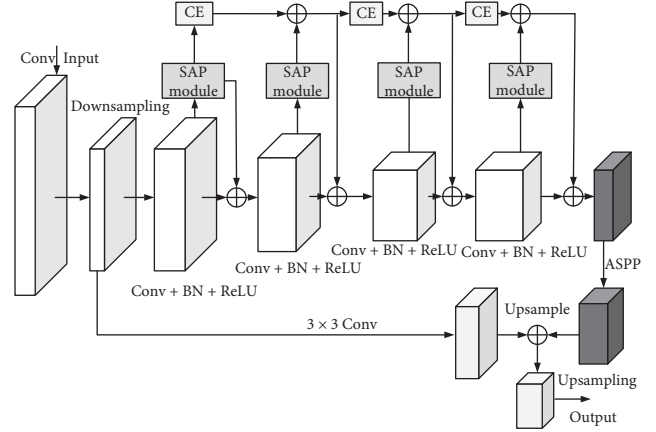


FIGURE 5: The proposed network structure.

TABLE 1: Experimental environment parameters.

Parameter	Configuration
Operating system	Ubuntu 16.04
CPU	i7-9700Ti
GPU	RTX 2080Ti
Internal storage	12G
CUDA	10.1
Python	3.6
TensorFlow	1.3.0

5×10^{-3} multiplied by $\sqrt{a^2 + b^2}$. Among them, $\text{power} = 0.99$ and $\text{iter}_{\text{current}}$ are the current number of iterations, and iter_{max} is the maximum number of iterations in the training process. The number of experimental training is 10,000, the image pixel size in the experiment is set to 512×512 , and the training batch size is $2 \times 8 = 16$. At the same time, during training, data enhancement measures such as random cropping, horizontal flipping, vertical flipping, and random sample scrambling are adopted. In addition, during the evaluation, the image will be scaled at multiple scales, with a zoom ratio of 0.55–1.55.

3.2. Experimental Data. The data used in the experiment come from the 2018 Data Science Bowl, which is manually labeled by professional doctors, and contain 670 pairs of original images with 9 resolutions and annotated segmented images of each nucleus, as shown in Figure 6.

The original image in the data set is shown in Figure 6(a). Each original image corresponds to multiple segmented images of a marked nucleus, that is, an original picture usually contains multiple nuclei, and the annotated images of multiple nuclei are merged as shown in Figure 6(b). When collecting raw data, different acquisition methods, different magnification magnifications, and different cell presentation methods are used. Moreover, the cell types collected are inconsistent. This results in the cell images in the data set with different morphology and distinct light and dark. The model needs to have a strong generalization ability and be able to adapt to a variety of different situations.

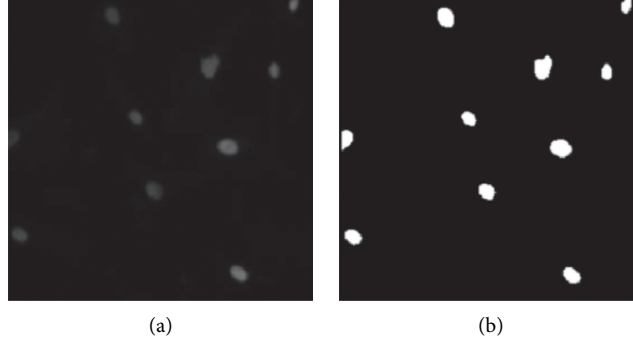


FIGURE 6: Nucleus image segmentation data: (a) original image and (b) annotated segmentation image.

3.2.1. Image Preprocessing. Due to the influence of various factors in the original data collection process, there is a large imaging difference in the cell images in the data set, which will affect the image segmentation. Therefore, it is necessary to preprocess it before segmentation. First of all, most of the pictures in the data set have a resolution of 512×512 . Therefore, the picture resolution is unified to 512×512 . Then, most of the data sets are grayscale images, and a few are color images. To improve the network processing speed, it is necessary to change the color images to gray scale.

At the same time, the contrast between the image of some nuclei in the data set and the background is small, which may make it difficult for the segmentation method to distinguish the nucleus from the background. Therefore, the data set needs to be preprocessed for histogram equalization. After the histogram equalization process, the gray value of cells and the background has a significant difference in the image, which helps the network extract more features.

In addition, in the process of collecting images, various noises are often interfered with and contaminated, resulting in a decrease in the signal-to-noise ratio of the image, and the edges of cells and the background become blurred. To improve the signal-to-noise ratio of the image, it is usually necessary to preprocess the image by filtering. In the experiment, a Gaussian smoothing filter is used to preprocess the image, and the filter $G(x, y)$ is calculated as follows:

$$G(x, y) = \frac{1}{2\pi\mu^2} e^{-\frac{(x-l-1)^2 + (y-l-1)^2}{2\mu^2}}, \quad (7)$$

where μ is the standard deviation of gray value, and l is the dimension of Gaussian convolution kernel.

3.2.2. Image Enhancement. To overcome the overfitting phenomenon of CNN, random shearing, flipping, gray perturbation, and shape perturbation are used in the experiment. Gray perturbation can transform each pixel in a small range. The gray value of the CT image is multiplied by a random number $[0.80 \sim 1.20]$, and a random number $[-0.20 \sim 0.20]$ is added. The grayscale perturbation in the training set can improve the stability of the network, thereby improving the performance of the prediction set network.

The CT and the contour images are deformed by affine transformation to form shape disturbance. The deformation

method is to obtain first the coordinates of the 3 vertices (upper left, upper right, and lower left). Then each point moves randomly, and the range of random movement is the image length. Finally, an affine transformation is performed on the entire image.

3.3. Evaluation Index. In the experiment, pixel accuracy (PA), mean pixel accuracy (MPA), and MIoU are used as indicators to evaluate the performance of the proposed network. Suppose N_{pq} represents the correct number of divisions; N_{pq} indicates that the original pixel belongs to the p category but is divided into the number of q categories; N_{qp} represents the number of pixels that originally belonged to the q category but were divided into p categories. There are $\tau + 1$ categories (including τ categories and an empty category or background category).

PA is the simplest accuracy measure for semantic segmentation, which represents the proportion of correctly marked pixels to the total pixels. The calculation is as follows:

$$PA = \frac{\sum_{p=0}^k N_{pp}}{\sum_{p=0}^k \sum_{q=0}^k N_{pq}}. \quad (8)$$

MPA calculates the proportion of pixels that are correctly segmented in each class and then finds the average of all classes. The calculation is as follows:

$$MPA = \frac{1}{\tau + 1} \sum_{p=0}^{\tau} \frac{N_{pp}}{\sum_{q=0}^{\tau} N_{pq}}. \quad (9)$$

MIoU calculates the ratio of intersection and union of two sets. The pixel intersection ratio is calculated within each pixel category, and then the average is calculated as follows:

$$MIoU = \frac{1}{\tau + 1} \sum_{p=0}^{\tau} \frac{N_{pp}}{\sum_{q=0}^{\tau} N_{pq} + \sum_{q=0}^{\tau} N_{qp} - N_{pp}}. \quad (10)$$

MIoU is highly representative, efficient, and concise and has become the current general image segmentation evaluation index. Therefore, MIoU is used as the main evaluation index of the experiment.

3.4. Training Process. When training the network, the input image undergoes local response normalization before the

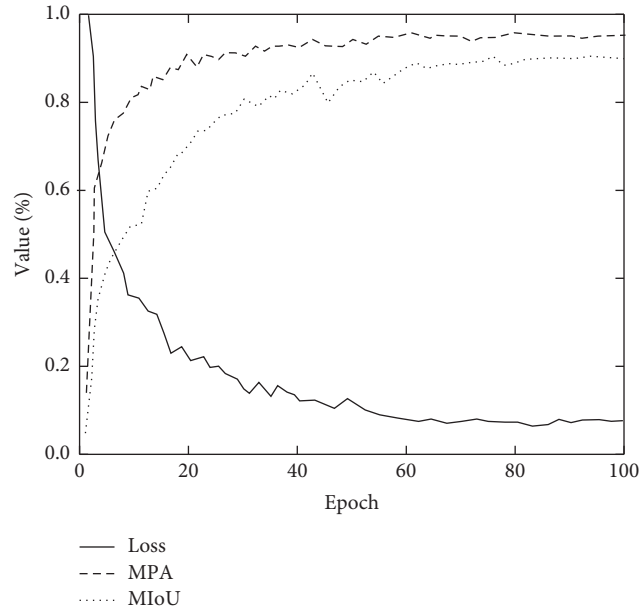


FIGURE 7: Curves of MIoU, MPA, and the loss.

first layer of convolution. The objective loss function is optimized by using the Adam algorithm with an initial learning rate of 0.005 and iterated until the loss function converges. The weight decay is 0.0001, and the number of iterations is set to 10,000. In the training process, the training data set is randomly shuffled, and then the batch size is set to 20. Due to the large deviation of the number of pixels in each category in the data set, the median frequency equalization method is used to balance between classes.

In the experiment, the proposed network is iteratively trained for 100 epochs on the data set. The changes in MIoU, MPA, and the loss of the validation set during the training process are shown in Figure 7.

MIoU reached 83% when the network was trained to the 50th epoch. In the subsequent training, MIoU stabilized, reaching 90%. MPA reached 94% when the network iterated to 60 epochs. Loss is reduced to about 7% when the network is iterated to 30 epochs. After 100-epoch iterations, the model basically converged, and the loss was reduced to below 4%.

3.5. Comparison of Technical Effects. After the nucleus image segmentation network converges, it is used to segment the images on the test set, and its segmentation effect is evaluated. In order to demonstrate the segmentation performance of the proposed technology, it is compared with references [14, 21]. The segmentation result of the nucleus image is shown in Figure 8.

It can be seen from Figure 8 that reference [14] uses differential evolution improved multi-level threshold to achieve image segmentation. The segmentation of the central part of the nucleus image is very good, but the segmentation of the edge details and smaller nuclei is not very good.

Moreover, there are certain over- and undersegmentation. Reference [21] proposed an automatic evolution model to achieve image segmentation. Similarly, the central part of the nucleus image is segmented well, and the phenomenon of over- and undersegmentation is relatively reduced. However, the result of manual annotation is relatively rough, and the ability to segment a smaller nucleus is poor. The proposed technology integrates the improved attention module into the U-Net network and has a better effect than the original U-Net for edge detail and smaller nucleus segmentation. The phenomenon of over- and undersegmentation is also relatively reduced, which is closer to the result of manual labeling. This proves that the proposed technology has ideal segmentation capabilities.

To quantitatively analyze the performance of the proposed technique, experiments are carried out on the data of nucleus image segmentation. Three indexes of PA, MPA, and MIoU are used to evaluate its segmentation performance with references [14, 21]. All test set data are used in the experiment to calculate the difference between the results of each technology segmentation and the manual segmentation standard, so as to obtain the evaluation results, as shown in Table 2.

The segmentation performance of the proposed technology is the best, and its MIoU reaches 85.02%. Because the proposed technology uses the most widely used U-Net network in the medical field and, at the same time, uses an improved attention module, the segmentation accuracy is further improved. Reference [14] realizes image segmentation based on a multi-level threshold improved by differential evolution. The optimal solution of the allocation population is generated by measuring the quality evaluation of the candidate solutions. The segmentation effect is affected by the choice of the larger optimal solution. Therefore,

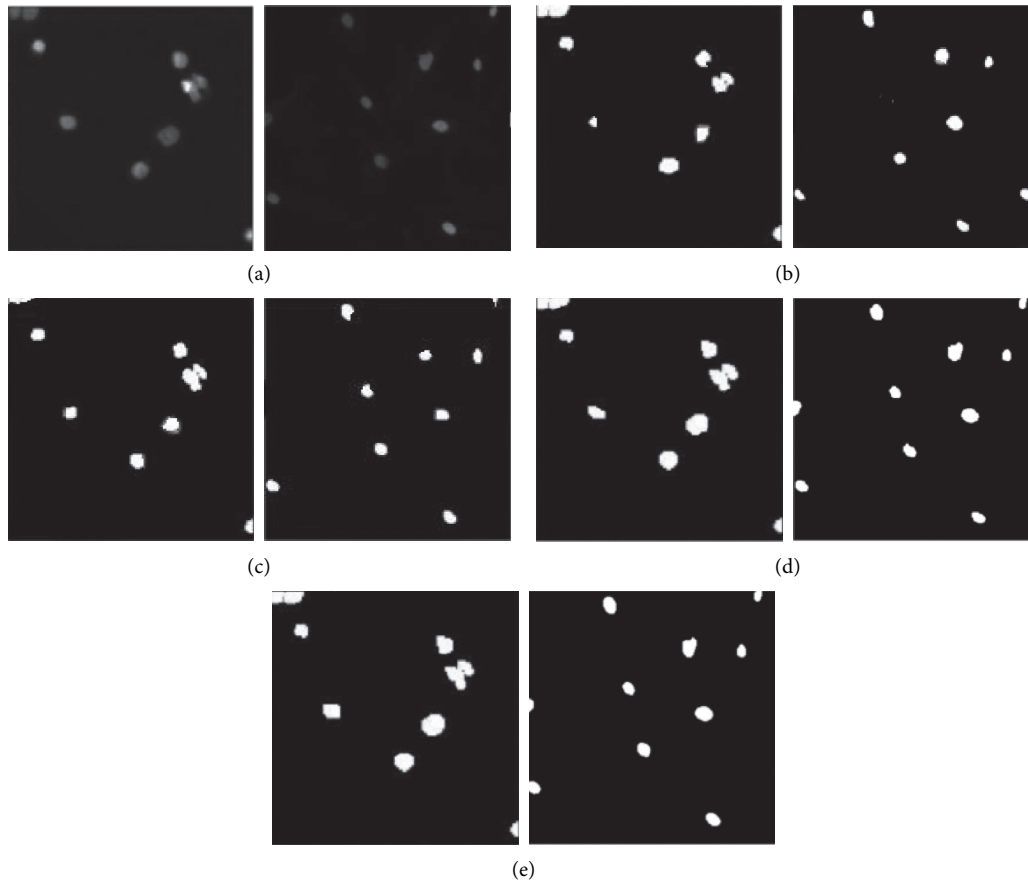


FIGURE 8: Comparison of nucleus image segmentation results: (a) original image, (b) image from reference [14], (c) image from reference [21], (d) proposed technology, and (e) manual marking.

TABLE 2: Performance comparison of three image segmentation networks.

%	Ref. [14]	Ref. [21]	Proposed technology
PA	82.54	85.68	89.97
MPA	84.17	87.39	91.35
MIoU	75.69	80.75	85.02

the overall performance is poor; MPA is only 84.17%. Reference [21] uses the evolutionary computing power potential of dense and residual blocks to propose an automatic evolution model, which has better image segmentation results. However, there are still some shortcomings in the image segmentation of a complex nucleus. Compared with the proposed technology, its MIoU is reduced by 4.27%.

Briefly, by comparing the experimental results, we can see that from the traditional segmentation technology in reference [14] to the deep learning algorithm in reference [21] and then to the improved U-Net network of the proposed technology, the segmentation effect and robustness of the nucleus are getting better and better. The segmentation effect of edge details and the smaller nucleus is also getting better and better. The experimental results show that the feature fusion method of skip connection and feature splicing in the U-Net network significantly improves the effect of U-Net image segmentation. Moreover, the

integration of the improved attention module has also significantly improved the accuracy of its image segmentation.

4. Conclusion

Traditional image segmentation algorithms generally need to extract some features manually in advance, such as the edges, corners, textures, and lines of the image. They have poor robustness and are easily affected by the environment. At the same time, the edges of the nucleus are more complicated, and the target is smaller. For this reason, a nucleus image segmentation technology based on the U-Net network is proposed. SAP and CE modules are used to improve the attention module, and the improved attention module is integrated into the U-Net network to segment the nucleus image, which further ensures the accuracy of segmentation. The 2018 Data Science Bowl data set is used on

the Python platform to demonstrate the proposed segmentation technology experimentally. The results show that it has better edge detail parts and smaller cell nuclei segmentation capabilities. PA, MPA, and MIoU are 89.97%, 91.35%, and 85.02%, respectively, which are better than other comparison techniques. It provides certain theoretical support for the high accuracy segmentation of the nucleus.

Because the offline method is used for image augmentation in the experiment, a larger storage space is required. In the following research, image augmentation can be added to the deep learning network to reduce the demand for storage space.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors wish to express their appreciation to the reviewers for their helpful suggestions that greatly improved the presentation of this paper. This work was supported by the Natural Science Foundation of Zhejiang Province (no. LY18F020002).

References

- [1] M. Tofighi, T. Guo, J. K. P. Vanamala, and V. Monga, "Prior information guided regularized deep learning for cell nucleus detection," *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2047–2058, 2019.
- [2] A. Veeramuthu, S. Meenakshi, and K. A. Kumar, "A neural network based deep learning approach for efficient segmentation of brain tumor medical image data," *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 1, pp. 1–8, 2019.
- [3] G. Wang, W. Li, M. Aertsen, J. Deprest, S. Ourselin, and T. Vercauteren, "Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks," *Neurocomputing*, vol. 338, no. 4, pp. 34–45, 2019.
- [4] M. Taheri, M. Rastgarpour, and A. Koochari, "A novel method for medical image segmentation based on convolutional neural networks with SGD optimization," *Iranian Journal of Electrical and Electronic Engineering*, vol. 9, no. 7, pp. 37–46, 2021.
- [5] U. S. L. U. Fatmatülzehra, C. Bass, and A. A. Bharath, "PERI-Net: a parameter efficient residual inception network for medical image segmentation," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 28, no. 4, pp. 2261–2277, 2020.
- [6] H. Zheng, Y. Zhang, L. Yang et al., "An annotation sparsification strategy for 3D medical image segmentation via representative selection and self-training," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 4, pp. 6925–6932, 2020.
- [7] D. A. Hasan and A. M. Abdulazeez, "A modified convolutional neural networks model for medical image segmentation," *Test Engineering and Management*, vol. 83, pp. 16798–16808, 2020.
- [8] H. Shang, S. Zhao, H. Du, J. Zhang, W. Xing, and H. Shen, "A new solution model for cardiac medical image segmentation," *Journal of Thoracic Disease*, vol. 12, no. 12, pp. 7298–7312, 2020.
- [9] T. Khin, K. Srujan Raju, G. R. Sinha, K. K. Khaing, and T. M. Kyi, "Review of optimization methods of medical image segmentation," *Proceedings of the Third International Conference on Computational Intelligence and Informatics*, vol. 1090, no. 1, pp. 213–218, 2020.
- [10] L. Fang, X. Wang, and L. Wang, "Multi-modal medical image segmentation based on vector-valued active contour models," *Information Sciences*, vol. 513, no. 4, pp. 504–518, 2020.
- [11] K. Li, L. Yu, S. Wang et al., "Towards cross-modality medical image segmentation with online mutual knowledge distillation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 1, pp. 775–783, 2020.
- [12] J. Junlong Cheng, S. Tian, L. Yu, and H. You, "Multi-attention mechanism medical image segmentation combined with word embedding technology," *Automatic Control and Computer Sciences*, vol. 54, no. 6, pp. 560–571, 2020.
- [13] E. Mabrouk, A. Ayman, Y. Raslan, and A.-R. Hedar, "Immune system programming for medical image segmentation," *Journal of Computational Science*, vol. 31, no. FEB, pp. 111–125, 2019.
- [14] H. Tarkhaneh, "An adaptive differential evolution algorithm to optimal multi-level thresholding for MRI brain image segmentation," *Expert Systems with Applications*, vol. 138, no. 6, pp. 112820–112829, 2019.
- [15] Z. Yang and A. Wu, "A non-revisiting quantum-behaved particle swarm optimization based multilevel thresholding for image segmentation," *Neural Computing and Applications*, vol. 32, no. 16, pp. 12011–12031, 2020.
- [16] F. F. Ximenes Vasconcelos, A. G. Medeiros, S. A. Peixoto, and P. P. Rebouças Filho, "Automatic skin lesions segmentation based on a new morphological approach via geodesic active contour," *Cognitive Systems Research*, vol. 55, no. 6, pp. 44–59, 2019.
- [17] D. Karimi and S. E. Salcudean, "Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 499–513, 2020.
- [18] A. Ratna Raju, S. Pabboju, and R. Rajeswara Rao, "Hybrid active contour model and deep belief network based approach for brain tumor segmentation and classification," *Sensor Review*, vol. 39, no. 4, pp. 473–487, 2019.
- [19] A. Geetha and N. Gomathi, "A robust grey wolf-based deep learning for brain tumour detection in MR images," *Bio-medical Engineering/Biomedizinische Technik*, vol. 65, no. 2, pp. 191–207, 2020.
- [20] L. Sun, W. Shao, M. Wang, D. Zhang, M. Liu et al., "High-order feature learning for multi-atlas based label fusion: application to brain segmentation with MRI," *IEEE Transactions on Image Processing*, vol. 29, no. 3, pp. 2702–2713, 2019.
- [21] T. Hassanzadeh, D. Essam, and R. Sarker, "An evolutionary DenseRes deep convolutional neural network for medical image segmentation," *IEEE Access*, vol. 8, pp. 212298–212314, 2020.
- [22] N. Kadoya, "Deformable image registration and auto-segmentation for various medical imaging types," *Igaku Butsuri*:

- Nihon Igaku Butsuri Gakkai Kikanshi = Japanese Journal of Medical Physics: An Official Journal of Japan Society of Medical Physics*, vol. 39, no. 1, pp. 12–19, 2019.
- [23] K. K. Jha and H. S. Dutta, “Nucleus and cytoplasm-based segmentation and actor-critic neural network for acute lymphocytic leukaemia detection in single cell blood smear images,” *Medical & Biological Engineering & Computing*, vol. 58, no. 8, pp. 1–16, 2019.
- [24] T. Wang, J. Huang, D. Zheng, and Y. He, “Nucleus segmentation of cervical cytology images based on depth information,” *IEEE Access*, vol. 8, no. 4, pp. 75846–75859, 2020.
- [25] J. C. Caicedo, A. Goodman, K. W. Karhohs et al., “Nucleus segmentation across imaging experiments: the 2018 data science bowl,” *Nature Methods*, vol. 16, no. 12, pp. 1247–1253, 2019.
- [26] T. Wan, S. Xu, C. Sang, Y. Jin, and Z. Qin, “Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks,” *Neurocomputing*, vol. 365, no. 6, pp. 157–170, 2019.
- [27] H. Zhang, H. Zhu, and X. Ling, “Polar coordinate sampling-based segmentation of overlapping cervical cells using attention U-Net and random walk,” *Neurocomputing*, vol. 383, no. 3, pp. 212–223, 2020.
- [28] J. Song, L. Xiao, M. Molaei, and Z. Lian, “Multi-layer boosting sparse convolutional model for generalized nuclear segmentation from histopathology images,” *Knowledge-Based Systems*, vol. 176, no. 15, pp. 40–53, 2019.
- [29] N. Ramesh and T. Tasdizen, “Cell segmentation using a similarity interface with a multi-task convolutional neural network,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1457–1468, 2019.
- [30] A. A. Mahmoud, W. El-Shafai, T. E. Taha et al., “An efficient segmentation technique for different medical image modalities,” *Menoufia Journal of Electronic Engineering Research*, vol. 30, no. 1, pp. 22–28, 2021.
- [31] M. Moirangthem and T. R. Singh, “Brain tumor detection through content-based medical image retrieval using ROI segmentation with Harmony search optimization,” *Journal of Green Engineering*, vol. 10, no. 10, pp. 8939–8969, 2020.
- [32] A. F. A. Fadzil, N. E. A. Khalid, and S. Ibrahim, “Amplification of pixels in medical image data for segmentation via deep learning object-oriented approach,” *International Journal of Advanced Technology and Engineering Exploration*, vol. 8, no. 74, pp. 82–90, 2021.
- [33] U. Reddy, P. Dhanalakshmi, and P. Reddy, “Image segmentation technique using SVM classifier for detection of medical disorders,” *Ingénierie des systèmes d’information*, vol. 24, no. 2, pp. 173–176, 2019.
- [34] E. K. Wang, C.-M. Chen, M. M. Hassan, and A. Almogren, “A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain,” *Future Generation Computer Systems*, vol. 108, no. 4, pp. 135–144, 2020.
- [35] D. Li, N. Deng, and X. Chen, “Level set medical image segmentation aided by cooperative quantum particle optimization with Lévy flights,” *Vibroengineering PROCEDIA*, vol. 28, no. 10, pp. 93–98, 2019.
- [36] Y. Niu and J. Cao, “Local difference-based active contour model for medical image segmentation and bias correction,” *IET Image Processing*, vol. 13, no. 10, pp. 1755–1762, 2019.

Research Article

Decisions on the Orderliness and Collaborative Operation Mechanism of Each Subsystem under the Organizational Model of the Internet of Things

Lingyi Cai  and Wei Liu 

School of Computer Science, Hubei University of Technology, Wuhan, Hubei 430068, China

Correspondence should be addressed to Wei Liu; 1910300628@hbut.edu.cn

Received 29 April 2021; Revised 25 May 2021; Accepted 1 June 2021; Published 10 June 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Lingyi Cai and Wei Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper analyzes the overall organic network system integrated by subnetworks, analyzes the network evolution process of the overall system of the Internet of Things (IoT) organization mode network, and analyzes the network through the network scale measurement, the node growth measurement, the node distribution measurement, and the node association measurement. We analyze the coordination mode of hardware resources, information resources, financial resources, human resources, organizations, and customer resources from the perspective of resource elements. And we analyze the vertical and horizontal work collaboration mode from the perspective of work activities. From the perspective of equity, two types of equity-based synergy and four types of contract-based synergy are proposed. It is proposed that the network stage of the IoT organization model is different from the previous stage of the important collaboration model, the platform-based collaborative model, which is analyzed from the coordination manager software platform and the IoT organization model infrastructure hardware platform. We use relevant data to construct a SIM model to measure the synergy of the development of the IoT industry system to quantify the overall collaborative development status of the IoT industry system. At the same time, the gray correlation analysis model GM (1, N) is used to correlate the complex system subsystems. The research results show that the development of the IoT industry has gone through a U-shaped development process from the uncoordinated stage to the coordinated development of the industrial system. Each subsystem has self-development capabilities and exhibits different interaction relationships with each other.

1. Introduction

The Internet of Things (IoT) is based on communication networks such as the Internet and mobile communication networks [1]. According to the application needs of different industries, it uses the perception, communication, and computing capabilities of agents to automatically obtain real-world information and connect all physical objects with independent addressing capabilities, so as to build an intelligent information service system that connects people and things and things and things. The IoT in specific application areas (smart agriculture, smart cities, etc.) still mainly stays in the links of environmental information perception, data transmission, and monitoring [2]. Although sensor technology is used, most of the collected data

are used for display or statistical analysis, and there is no information linkage with related control equipment. In fact, scientific decision-making and intelligent control are not realized. The more common applications are perception without decision-making or decision-making without control. These applications are single-step applications, which have not yet constituted the “closed loop” of the entire link of “sensing-decision-control,” which greatly weakens the IoT technology [3].

Under the new economic form, there are new development forms of the IoT organization model and development needs of the IoT organization model. Scholars supplement and improve the theoretical system of the organization model of the IoT, combine the current economic development status and trends, make new interpretations

and more in-depth studies, and propose the network system planning and operation methods of the organization model of the IoT on this basis [4, 5]. It is of great theoretical significance to enrich and develop the theoretical system of the subject of the organizational model of the IoT. The network construction optimization model studied in this paper, as well as the network coordination operation mode and coordination mechanism, realizes the dynamic sharing of information and resources of the main body of each dimension in the network, the coordination and integration of operation processes, and better specialization and integration of services for regional economic industries. It is expected to provide scientific methods and basis for the Chinese government and enterprises in the IoT organization model planning, promote the development of the IoT organization model industry cluster, optimize the overall operating performance of the network system, reduce the overall operating cost of the IoT organization model network system, optimize IoT organization model network resource allocation, reduce process links and time-consuming, improve the degree of cooperation and coordination of the IoT organization model network system, promote the increase of the value of the IoT organization model, and have practical guidance for promoting the development of the IoT organization model industry significance [6].

Coordinating managers can exist stably and effectively promote collaborative management work and must pass a series of institutional guarantees and government support. First, the establishment of a coordinated management organization needs to be led by relevant government departments, with the participation of key IoT organizations and park management committees. Then, it is responsible for formulating and implementing the system's task coordination execution mechanism, information sharing mechanism, collaborative incentive mechanism, and collaborative behavior. The supervision data are obtained by means of information, and the reward and punishment results and early warning information are output through the established mechanism model to assist in the implementation of the mechanism. Based on the measurement of the degree and status of the development of the IoT industrial system, a gray dynamic correlation analysis model of the synergy among the various subsystems within the IoT industrial system, namely, the GM (1, N) model, is constructed according to the gray system theory. At present, each subsystem has the ability to develop itself. Among them, the fusion subsystem and the technological innovation subsystem have a one-way weak synergy relationship.

1.1. Related Work. The IoT is a very complex heterogeneous network, which builds a bridge between the physical world and the virtual world [7]. The ultimate goal of all IoT applications is to provide seamless services without human intervention. The IoT is considered to be the next logical revolution, which can provide a wide range of services in smart cities, manufacturing, smart agriculture, smart healthcare, and smart homes [8, 9]. Autonomous IoT systems are very important, but there are still many challenges

to be solved. Stankovic mentioned in the inaugural IEEE "Journal of the IoT" that the IoT should be an important infrastructure that can run multiple applications and services and has integration capabilities [10]. Scholars proposed a kind of cognitive IoT, including how to deal with heterogeneous data and high-dimensional data of the IoT, the discovery of the semantics of the IoT, and the interpretation of related technical issues such as the intelligent decision-making of the IoT and in-depth research on the intelligent services of the IoT. The IoT has quietly entered multiple smart industries. In today's advocacy of energy conservation and environmental protection, the optimization of the resource allocation of the IoT has become an urgent problem to be solved [11]. There are many artificial intelligence algorithms for resource scheduling, but most of them are developed for specific applications and have certain limitations, which are not suitable for solving the IoT service problems studied in this paper [12]. As the IoT is a service, this paper will treat its entire layout as a service system. The solution of the resource optimization allocation problem of service-oriented networked collaborative equipment is a very complex problem, which belongs to the typical NP-hard combinatorial optimization problem [13]. The question studied in this paper is how to reduce resource consumption and shorten service time, in other words, how to select multiple optimal services from a large set of candidate services to meet the above goals. This will be a challenging multiobjective optimization problem.

Many researchers try to solve the multiobjective service selection problem in Web services [14]. Scholars proposed the first approximation of multiobjective quality-driven service selection [15]. Related scholars have carried out research on multiobjective optimization of service quality. They introduced the Pareto set model for service composition with service quality perception. Related scholars have proposed a method to support decision-makers to use clustering to find robust and QoS-optimized service combinations [16]. In our previous research work, we implemented an adaptive Web service composition inspired by the neuroendocrine-immune system. However, the above-mentioned research work mainly focuses on Web service composition based on service quality. IoT services is different from Web services in that they are characterized by large-scale, heterogeneous, unreliable, and dynamic nature. An important challenge that needs to be solved in the field of IoT service composition (also known as service integration) is to develop efficient service selection algorithms in order to optimize management of energy consumption and service quality. In a large-scale IoT environment consisting of thousands of distributed entities, this issue becomes critical. Relevant scholars use probabilistic discovery methods to efficiently find an approximate service set that satisfies a given request to meet the goals of rapid service discovery and minimize resource consumption [17]. Related scholars pointed out that the IoT is a paradigm, in which physical things in the real world can be connected to the Internet and provide services through additional computing devices. Researchers have proposed a three-tier service quality scheduling model for the service-oriented IoT. Perception as

a service model is expected to be built on IoT infrastructure and services. Then, the service is allocated to the interface with heterogeneous resources, and the optimal solution is generated for this NP-hard problem [18].

The large-scale IoT environment is composed of thousands of distributed entities. Once a service request is detected, it is necessary to immediately calculate and select multiple (request-service) services with the optimal combination of request-service pair from a large number of candidate sets to provide the service. The IoT should not only provide services for dynamic concurrent requests but also reduce energy consumption, reduce service time, and improve information accuracy. In addition, service availability, bandwidth allocation, and reliability should also be considered, especially in resource-constrained environments [19]. Therefore, service selection is a multiobjective optimization problem. However, the development of effective service selection algorithms is still challenging and has not been extensively studied. Recent research results, such as the use of novel decomposition strategies, heuristic caching strategies, and heuristic multitask scheduling algorithms, are applied to the IoT service selection problem, but they have not been in-depth from the perspective of multiobjective optimization [20]. The optimization of IoT service portfolio based on resource and QoS awareness and the optimization of IoT service portfolio based on energy consumption and QoS achieve dual-objective optimization, but the dynamic characteristics of the service are not considered [21, 22]. Therefore, the purpose of this paper is to achieve global combination and multiobjective optimization of ubiquitous services in a dynamic IoT environment.

2. Core Technology for Decision-Making Collaboration

2.1. The Process of Decision-Making Coordination. The ultimate goal of autonomous decision-making is to better conduct penetration testing, that is, detection, intrusion, and information theft of the target system. After obtaining important information, it is necessary to use the same channel and encryption method as legal data for information transmission through controlled distributed clients and disguised applications, concealing itself while avoiding security audits and anomaly detection mechanisms.

Effective and reasonable attack decision-making will realize the migration of multiple objects. Therefore, autonomous decision-making needs to have a clear recipient object and target (including attack range, target effective value, attack time limit, and damage degree) at specific stages, the system type, defense mechanism, and security equipment deployed are comprehensively planned, and the data collected in the information collection phase are dynamically adjusted during the simulation attack phase to achieve the optimal attack effect, transfer to the next object, and finally achieve the target system penetration effect.

This paper mainly involves the decision-making system and collaboration system, as well as the later development of the penetration test module. The whole working process is shown in Figure 1.

The whole system includes three stages: decision-making stage, collaboration stage, and penetration testing stage. The decision-making stage includes four main functional modules, namely, authentication module, node selection module, task expression, and task segmentation. The collaboration stage is divided into functional modules such as task distribution, idling, communication, and exit according to the inherent needs of the collaborative system and the process of collaborative processing.

2.2. Task Segmentation. In order to increase the effectiveness of decision-making, the task constraint conditions set up are expressed with the constrained object as the center. For example, in the decision-making authentication stage, there are restrictions on authentication information, including key/certificate information, version information, and so on. The target object usually includes constraints such as operating system, network environment, and application environment. It is necessary to divide the task into basic executable tasks. When the task is a compound task, the “task segmentation and scheduling” must be completed before the task is executed.

In order to achieve the balance of the system and simplify the algorithm, it mainly examines the hardware conditions of the node such as CPU, disk, and memory, as well as the network information and operating system version of the node. These characteristics are combined to make node selection. Different nodes are selected for the decision tasks of different target hosts. For example, when sniffing and scanning the hosts in the same local area network, it requires less resources, but higher requirements are placed on the network of the node. At this time, the network information and other parameters of the node are mainly used as the basis for selection. When the task becomes a password cracking that requires more hardware resources such as CPU and memory, the node’s choice of focus is converted to the node’s hardware as the main reference basis. This is because password cracking attacks often require a lot of local resources. Of course, network resources have also been improved. Therefore, when designing the node selection algorithm, we need to combine tasks for node selection. For this, you can refer to the task type after task segmentation in the previous section to set the corresponding selection coefficient.

2.3. Cooperative Communication. The design of the cooperative communication module is for the communication between cooperative nodes and between cooperative nodes and cooperative controllers. The protocol interaction method involves the communication between the two ends of the C/S. Channel is used to maintain the “channel” information, including the SOCKET (abstract description) of the communication and the option information of the SOCKET. The design is based on the support of multiple operating system platforms. By maintaining the attribute information corresponding to the socket handle, using Channel can facilitate other modules to modify these attributes and increase these interfaces. After combining the

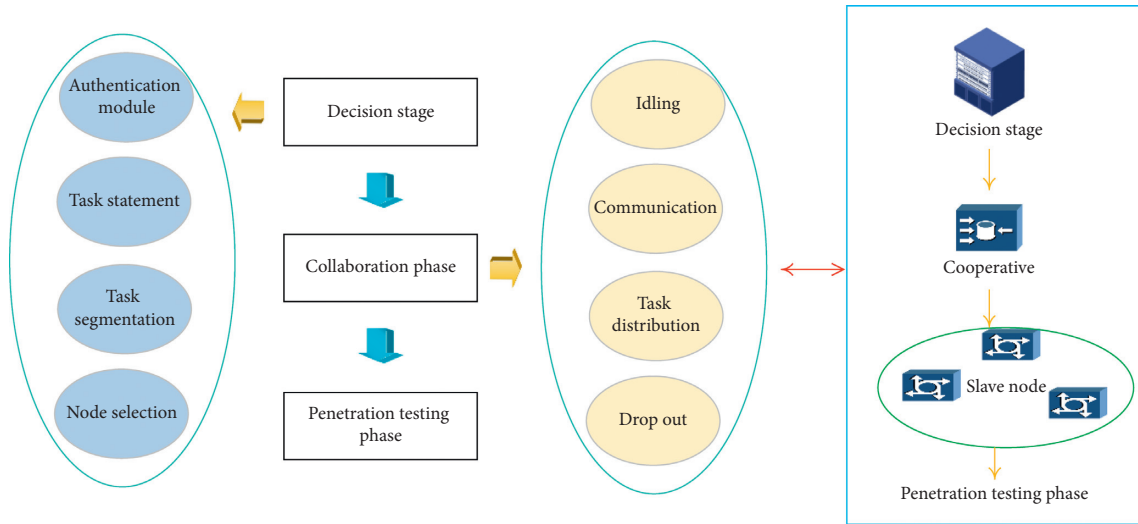


FIGURE 1: System working process.

data management, the corresponding implementation results should meet the model diagram shown in Figure 2.

3. Decisions on the Collaborative Operation Mechanism of the Network System of the Multidimensional IoT Organization Model

3.1. The Principle and Process of the Orderly Cooperative Operation Mechanism of Each Subsystem. The collaborative operation of the IoT organization mode network is a series of IoT organization mode activities that are oriented to customer needs and revolve around the customer's IoT organization mode service tasks. In the IoT organization model network system, through the coordination management platform, m customer tasks are accepted and decomposed into n IoT organization model operations, and the paths of these operations are optimized and integrated. You collect k IoT organization model service providers that can provide service resources through the network IoT organization model platform and evaluate and match operation resources. From the overall perspective of the system, you call the resources in the network system, select the optimal configuration of the resources, and work together to complete the tasks of the IoT organization mode by the division of labor selected by these preferred resources. This collaborative operation mechanism is based on customer needs. The mechanism enables the resource distribution of the IoT organization model service provider to be guided by the overall operation requirements of the system, continuously optimizing the distribution and migration and realizing the optimization of the overall system operation.

The collaborative operation mechanism of the IoT organization model network is the core of all other collaborative mechanisms. Node members can carry out collaborative behaviors such as division of labor and cooperation in accordance with the rules of the collaborative operation mechanism and then can further implement the

incentive and supervision mechanism of collaborative behavior on this basis. The collaborative operation mechanism studies how to integrate operations and achieve economies of scale and distance economic effects and how to match resources to give full play to the respective advantages of the IoT organization model service providers, reduce vicious competition, and reduce the return load rate, thereby reducing the overall cost of the network. This process mainly includes the three processes of job grouping, job integration, and resource matching execution. The rules involved in the three stages include main job sorting rules, job integration rules, and resource matching selection rules, as shown in Figure 3.

In the node operation, the operations of the same location and the same type are integrated. The integration of transportation operations is more complicated. If the scale effect is not considered, each operation chooses to take the shortest path between the starting point and the end point, but it may lose the opportunity to integrate with other operations. It is not necessarily the overall best choice. In the same group of transportation operations, according to the evaluation of the main operations from high to low, the shortest paths of the operations are found in turn, and the nonmain operations are integrated into the path of the main operations in order to achieve economies of scale. The step is to list the shortest path of each job. The Dijkstra algorithm can be used, and the number of iterations is set to K ; that is, K shortest path solutions are listed. The second step is to integrate the operation path and determine which operations can be integrated into the main operation path. If the cost saved by the scale effect can offset the detour cost and transshipment cost, then integrate; if the increased cost cannot be offset, then do not integrate.

We match the integrated work according to the resources of the nodes and lines and find the resources in the path according to the previous path plan. According to the resource situation of the service provider, it can be completed by a single service provider or multiple service providers. This process is effective recombination of the

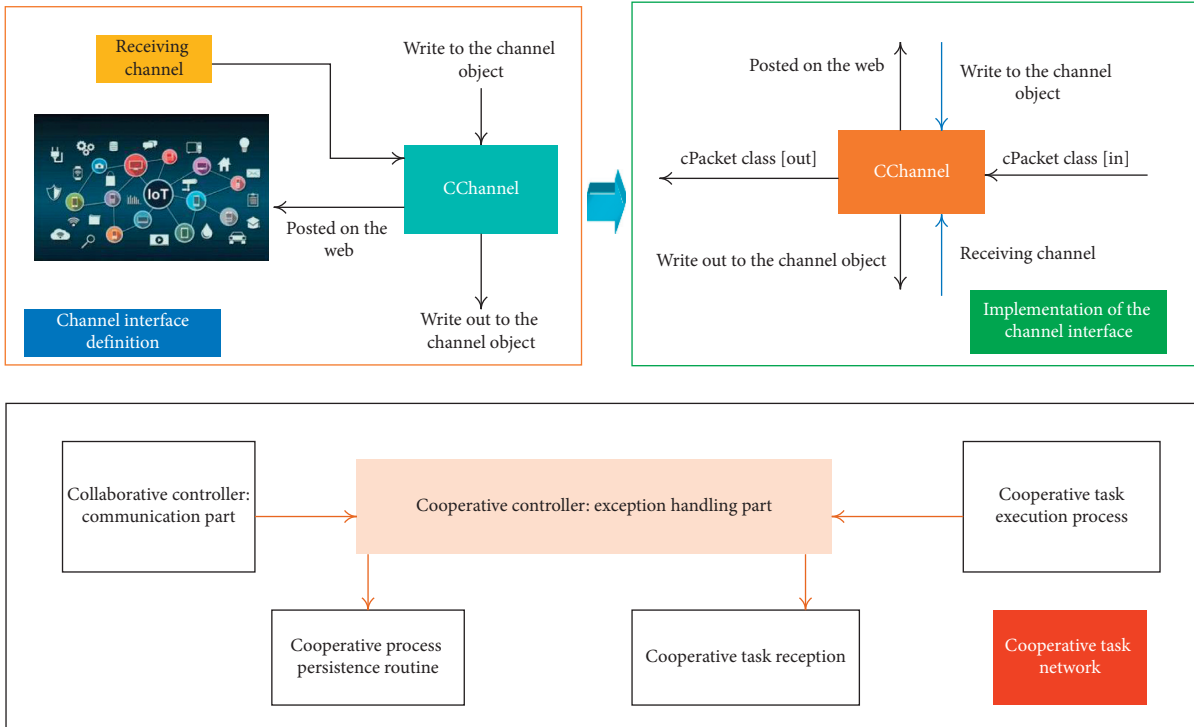


FIGURE 2: Schematic diagram of the model after the channel interface is implemented.

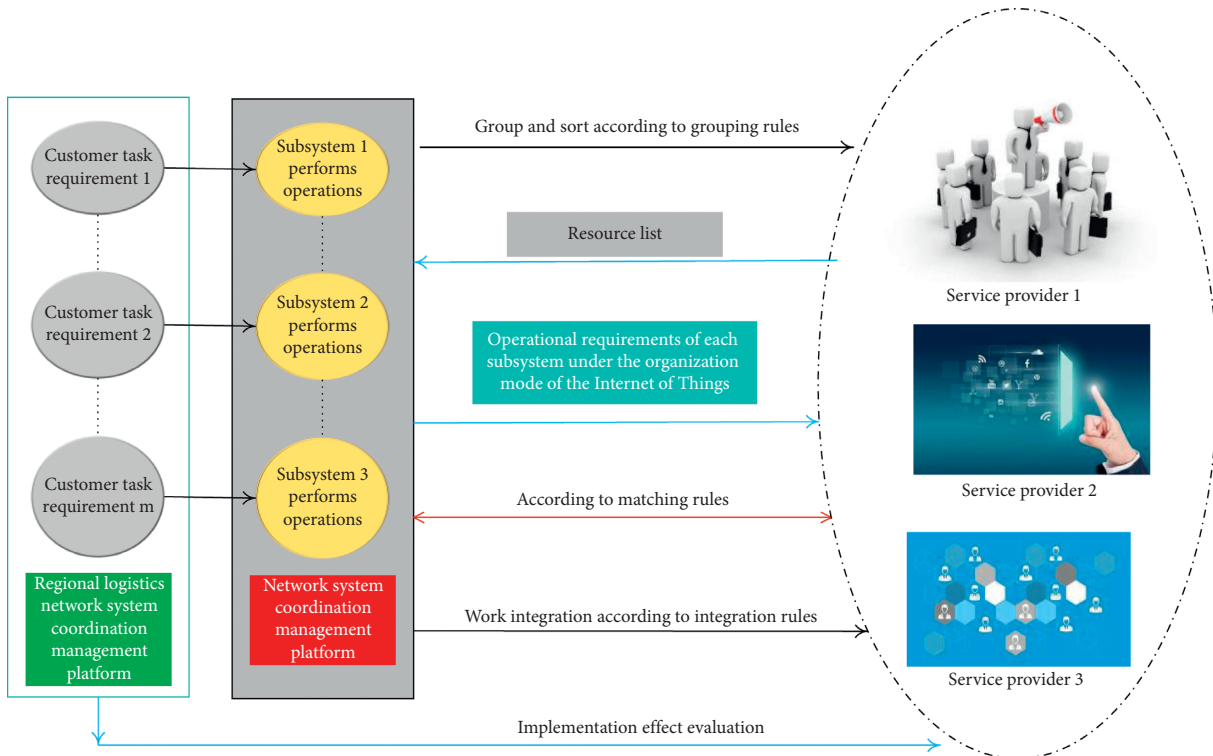


FIGURE 3: The collaborative work process of the IoT organizational model network.

service resources of the distributed IoT organization model. It is the redesign of the organization model network of the IoT organization model to quickly respond to market demand for its own process and organization.

According to the published service requirements, the resource search is performed according to the search conditions, and the information of the candidate companies that match the conditions is displayed in feedback. You give

priority to the resources and success history of the service provider and then judge whether the resources can complete the task. When the working capacity A_i of the preferred service provider is sufficient, the job T_i of the node or the line can be completed independently. When the existing resources of the service provider cannot meet the demand, the suboptimal service provider is selected in turn to subcontract to complete the operation. If the total resources of the service provider at this point (route) are still insufficient, you apply to call the idle resources of the service provider of other nodes (routes).

3.2. Establishment of an Orderly Collaborative Operation Mechanism Model for Each Subsystem. Suppose that the network graph of the IoT organization model is composed of points and edges, $G = (V, L)$, where V is the set of nodes, m is the total of points, and L is the set of edges. G becomes the abstraction of the nodes and traffic connections of the IoT organization model in the geographic map. Due to the limitation of transportation capacity of service providers, each service provider can only transport part of the road sections in the transportation network. Therefore, the network diagram of the IoT organization mode is a mixed road network diagram, and the side variable $Y(v_i, v_j) = S_k$ ($k = 1, 2, \dots, K$), where S_k represents the service provider, indicating that the edge (v_i, v_j) is the transportation route of the service provider S_k ($k = 1, 2, \dots, K$), and the point variable $B(v_i) = S_k$ ($k = 1, 2, \dots, K$), indicating that vertex v_i is the job node of the service provider S_k . R_{ij} is the number of service providers between points i and j . S_{ijk} is the k -th service provider between points i and j .

Suppose that there are N jobs in the network to form job set Z ; Z_n ($n, V_a, V_b, q_n, T_{\max}, C_{\max}, \text{type}$) means that the

starting point of the n th job Z_n is V_a , the end point is V_b , and the amount of work is q_n . For time requirements, C_{\max} is the cost budget of the customer; that is, the time for customer satisfaction cannot exceed T_{\max} , and the cost cannot exceed C_{\max} . Type represents the job type and is used for the group number of the job.

Suppose that there are a total of K service providers in the network to form a service provider set S . S_k means that the operation capacity of the k -th service provider S_k between the service range V_i to V_j is A_{ij} , the total amount of work undertaken is Q_{ij} , i, j, n the ability to undertake the service is evaluated as D_{ij} , the service price C_n represents the price required to complete the transportation operation of Z_n , C'_n represents the price required to complete the node operation Z_n , T_n represents the completion of the transportation operation, T'_n represents the time required to complete the node job Z_n , and type is the resource type of the service provider, which matches the type grouping in the job variable.

$$\begin{aligned} Q_{ijk} &= \prod_{n=0}^{N-1} \gamma_{nijk(Rn)} \bullet q_{n-1}, \\ C_n &= q_{n-1} \bullet L_{ij} \bullet C_{ij} \bullet Q_{ij}, \\ C'_n &= q_{n-1} \bullet C'_{ij} \bullet Q_{ij}. \end{aligned} \quad (1)$$

Among them, $C_{ij}(Q_{ij})$ is the price per unit of distance per unit of work, which is a descending function or a segmented discount function. The larger the total amount of work undertaken by the service provider S_k in this segment, the lower the price per unit of work, which reflects the effect of scale.

$$C_{ij}(Q_{ij}) = \begin{cases} \alpha_1 c_{ij}, & Q_{ij} > b, \\ \alpha_2 (1 - c_{ij}), & a < Q_{ij} < b, T_n = \frac{(L_{ij} - 1)}{|V_{ij}|}, T'_n = \frac{(q_n - 1)}{|V'_{ij}|}, \\ c_{ij}, & 1 < Q_{ij} < a, \end{cases} \quad (2)$$

3.3. The Process of Solving the Orderly Coordination Mechanism of Each Subsystem. The two-layer iterative method is used to solve the problem, and the solving process is shown in Figure 4. The solution is divided into two levels. The first level is job reorganization. First, the main job-based job boarding method is used for integration and optimization, so that resource matching is guided by system optimization, and the service provider selection is the second level problem. The suboptimization problem of the problem can be solved by using the neighborhood search method.

3.4. Establishment of an Incentive Mechanism Model for Orderly Coordinated Behavior of Each Subsystem. The idea of this mechanism model is mainly to further improve on the

basis of the price management mechanism model and carry out collaborative gain distribution. The synergy gain is the difference between the system's optimal return E_{total} and the noncooperative total return E_{total} . It is assumed that a part of the gain will be distributed to all members of the IoT organization model service provider participating in the collaboration, and the other part will be rewarded to the IoT organization model service provider with excellent service and active network construction behavior. That is to say, the income of the service provider is divided into the basic subcontracting income plus the reward income.

Suppose that there are K IoT organization model service providers in the IoT organization model network system, the annual subcontracting operation volume of the k th IoT

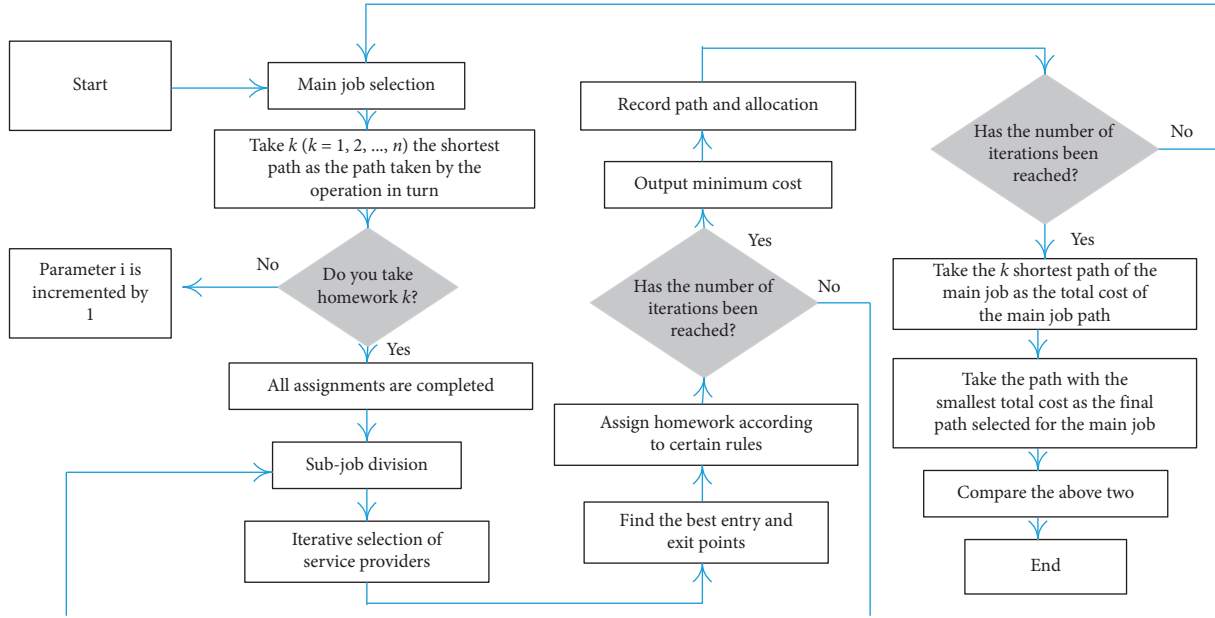


FIGURE 4: Overall solution process of orderly coordination mechanism of various subsystems.

organization model service provider is q_k , and the operation execution result is $U_k = (u_{k1}, u_{k2}, u_{k3})^T$. Among them, u_{k1} is the basic IoT organization mode operation assessment, including on-time arrival rate and customer service evaluation. u_{k2} is the information construction index of the organization model of the IoT, such as the timely rate of information online, the coding rate of equipment standardized information, and the investment in information construction such as EDI. u_{k3} is the coordination and cooperation index with other IoT organization model service providers, such as the completion rate of loading and unloading preparation and the timely rate of vehicle preparation; before the arrival of the downstream subcontractor's transport vehicle, we prepare for shipment. These three types of evaluation indicators are integrated into the 360-degree inspection method; that is, the company itself, customers, the coordinator of the IoT organization model platform, and the upstream and downstream cooperative companies jointly evaluate. Most of the evaluation information is automatically entered into the calculation method based on objective work conditions to reduce management costs brought by incentives. The subcontract of the service provider to improve the organization model of the IoT is

$$S_k(U_k) = \prod_{i=0}^2 q_k \cdot \alpha_{ki} |u_{ki} - u_{ki0}| \cdot p_{k-1}, \quad k = 1, 2, \dots, n. \quad (3)$$

Among them, the first item is incentive payment, and the second item is fixed standard payment. u_{ki0} is the up-to-standard value of each index, and α_{ki} is the incentive coefficient of each index, which is also the decision variable to be solved in this section. p_k is the fixed payment for completing the job task and meets the P_i interval range of the previous section; set

$$\begin{cases} u_{k1} = e_{k1} - |\varepsilon_{k1}|, \\ u_{k2} = e_{k2} - |\varepsilon_{k2}|, \\ u_{k3} = e_{k3} - |\varepsilon_{k3}|. \end{cases} \quad (4)$$

Among them, ε_{ki} ($i = 1, 2, 3$) represents the effort spent by the k -th IoT organization model service provider in the i -th type indicator. ε_{ki} ($i = 1, 2, 3$) represents the external objective factors that affect the completion of the i -th type index of the k -th IoT organization model service provider. The external objective factors that affect the operation index of the IoT organization model include weather conditions, traffic conditions, and natural disasters; external objective factors that affect information sharing include the service conditions of public telecommunication operators and the information network construction of the IoT organization model network platform.

Suppose that the benefits brought by the j -th IoT organization model service provider to the IoT organization model network system are

$$H_k(U_k) = \prod_{i=0}^2 q_{k-1} \cdot u_{ki} \cdot \beta_{ki}. \quad (5)$$

Among them, β_{ki} reflects the contribution of the i -th index to the system revenue. The first category of indicators is the completion of the organization model of the IoT, the second category of indicators is the status of information construction and sharing, and the third category of indicators is the status of coordination and cooperation.

Suppose that the coordination cost between the network platform coordinator of the IoT organization model and the k -th IoT organization model service provider is

$$C_k = \frac{1}{k} \cdot \prod_{i=0}^2 e_{ki}^2. \quad (6)$$

Therefore, the total benefits of all IoT organization model service providers in the IoT organization model network system to the coordinator are

$$G = \prod_{k=0}^{K-1} q_{k-1} \cdot |S_k + C_k - H_k|. \quad (7)$$

4. Evaluation and Analysis of the Synergy Degree of the IoT Industry System

4.1. Determination of Indicator Weights. The degree of influence of each subsystem in the evaluation index system on the overall system is not the same, which requires the corresponding weight to be assigned to each evaluation index. This paper uses the correlation matrix weighting method to determine the indicator weights. The correlation coefficient reflects the degree of interaction between indicators. The larger its absolute value, the closer the relationship between the indicators. The deeper the impact, the lower the converse. If the overall correlation between an indicator and other indicators in the evaluation system is relatively high, it means that this indicator has a greater impact on the development and changes of other indicators; that is, it has a greater role in the system. Assuming that the index system contains n indicators, the correlation matrix between the indicators is represented by R ; let

$$R_j = \prod_{i=0}^{n-1} (1 - r_{ij}), \quad i = 0, 1, \dots, n-1. \quad (8)$$

Then R_i represents the total impact of the i -th indicator on other $n-1$ indicators in the indicator system. The larger R_i , the greater the impact of the i -th indicator in the indicator system and the greater its weight. The corresponding index weight can be obtained by normalizing R_i ; namely,

$$\lambda_i = \frac{R_i}{\prod_{i=0}^{n-1} R_i}, \quad i = 0, 1, \dots, n-1. \quad (9)$$

The calculated weight of each evaluation index and the proportion of each index are shown in Figure 5.

4.2. Measurement and Evaluation of the Synergy Degree of the IoT Industry System. The coordination degree of each subsystem of the IoT industrial system and the coordination degree of the IoT industrial system are calculated. The weights of the four subsystems are all set to 0.25 when calculating the coordination degree of the composite system. The estimated results are shown in Figure 6.

Figure 7 shows the degree of synergy among the various subsystems of the IoT industry system. It can be seen from the figure that the coordination status of each subsystem is more complicated. The degree of synergy of the fusion subsystem is rising in volatility, and after reaching the peak, it shows a downward trend to the lowest point. The decline in the degree of synergy of the fusion mechanism shows that the industrial barriers have not been eliminated, and it is still hindering the coordinated development of the IoT industry.

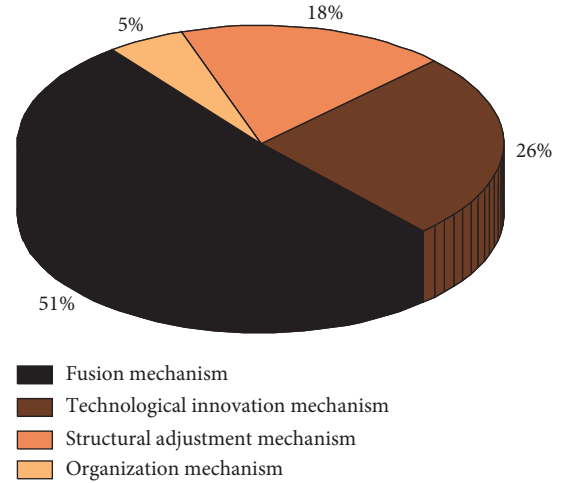


FIGURE 5: Proportion of the evaluation index of the synergy degree of the IoT industry system.

After that, it showed a trend of picking up again, indicating that industrial integration has gradually developed slowly after experiencing a process of disorderly development. The reason is that the IoT industry chain is long, and the differences between different industries have hindered product interaction to a certain extent. Effective communication between the information collected by different industries in this layer is different, so the available sensors are different. For example, the sensors used for soil monitoring and the sensors used for environmental monitoring are not universal. The same is true at the information processing level. The content they need to analyze and make decisions about is different from each other. In the current period of development, due to the diversified nature and strong customization of the industry chain including industries, it is difficult for IoT technology and industry growth to form a centralized leapfrog development. After reaching the peak point, the organizational mechanism synergy showed a downward trend, indicating that the organizational mechanism that can lead the development of the IoT industry has not yet been perfected.

The growth synergy degree of the IoT industry system is shown in Figure 8. From the perspective of the degree of synergy of the growth system of the IoT industry, the development and growth of the IoT industry have gone through a wave-shaped development process from the uncoordinated development of the current industrial system. IoT-related companies are generally small in scale and weaker in resisting risks and are more affected by the financial crisis. On the other hand, since the development of the IoT was proposed, it has brought opportunities and problems to the development of the IoT. With regard to the IoT, scientific research institutions and local governments have begun to rush to get a share of government investment. This has caused a serious situation of low-level redundant construction. The future development of the IoT has a broad space, but it is separated by many different application scenarios, forming a large number of small markets with limited capacity, which has caused a phenomenon that

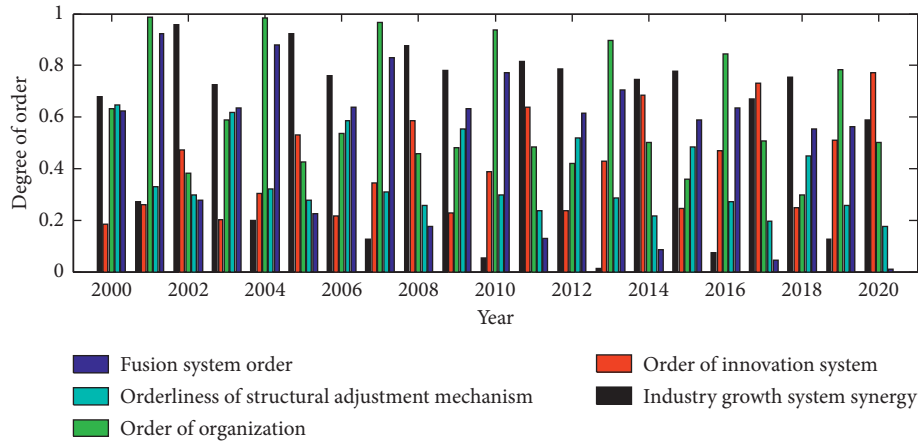


FIGURE 6: Estimation of the degree of synergy in the IoT industry system.

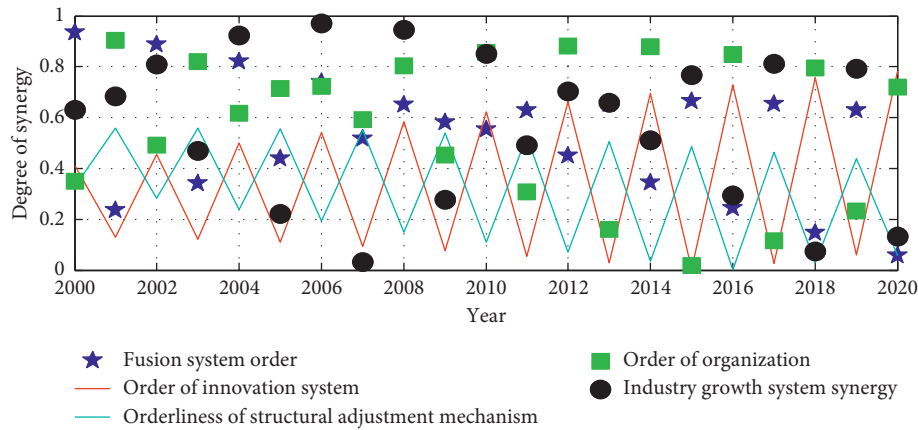


FIGURE 7: Changes in the degree of synergy among the various subsystems of the IoT industry.

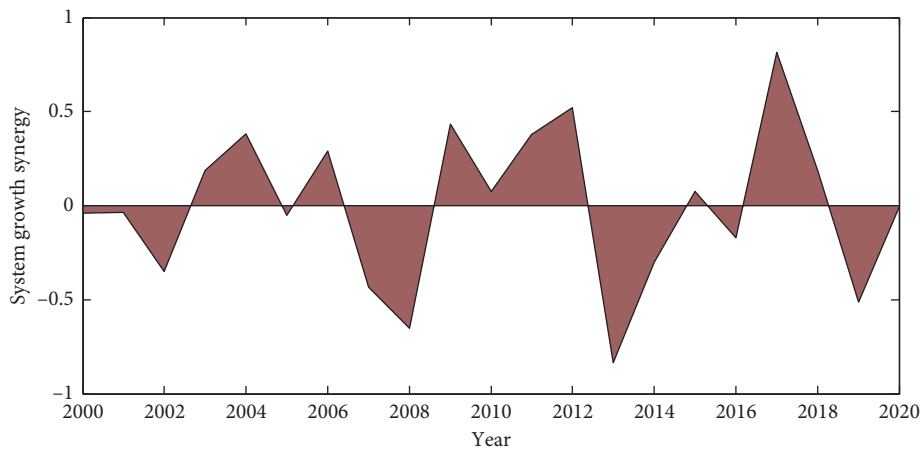


FIGURE 8: The degree of synergy in the growth of the IoT industry system.

diversified application scenarios limit the development of industrialization. This phenomenon is mainly manifested in two aspects. On the one hand, in the initial stage of the development of the IoT, the technology application and product development are not yet mature, and some of the IoT technologies are relatively high-end, and they lack

advantages in price, which cannot satisfy the market's demand for product functions. It is difficult to obtain product promotion and large-scale development due to the requirements of sex and price; on the other hand, some of the current technologies cannot be used to create value and function by realizing productization. Now they only stay at

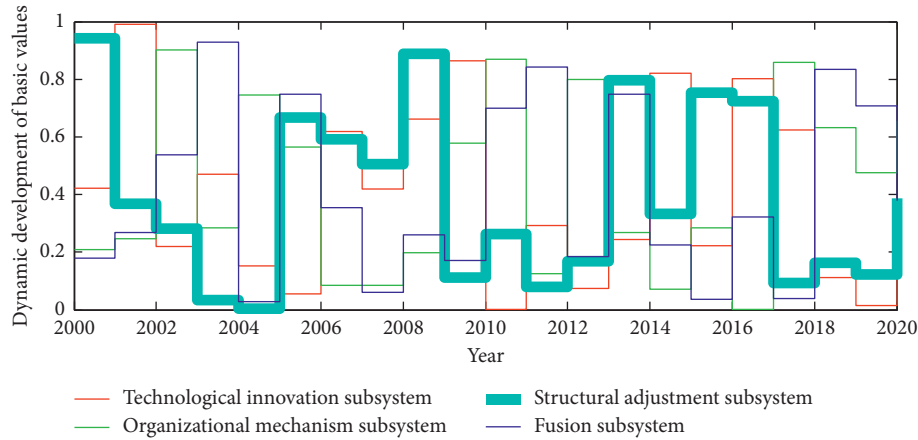


FIGURE 9: The basic data sequence of the dynamic development of each subsystem of the IoT industry.

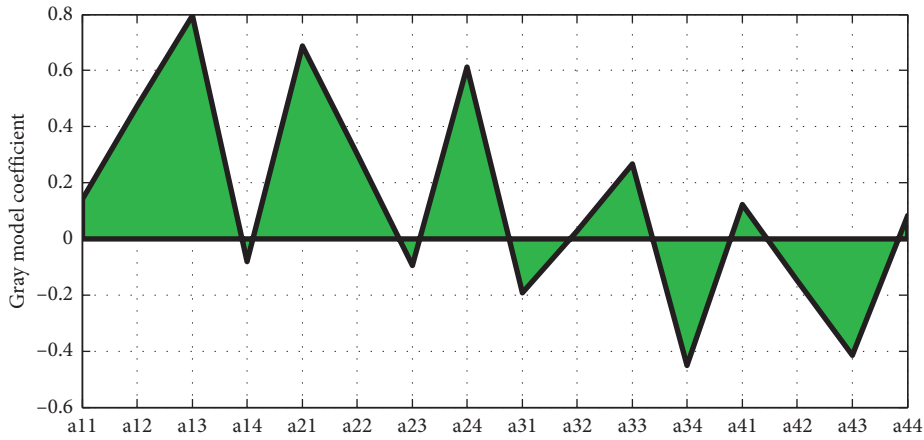


FIGURE 10: Gray model coefficients of various subsystems of the IoT industry.

the level of laboratory and simulation, which is difficult to form. Large-scale applications cannot promote the improvement of the industrialization level. For example, based on the application of sensors and sensor networks, the conflict between application scale and industrialization is very obvious. For example, the country does not have the ability to produce sensors for detecting soil and water quality. From the comparison of the order degree of the four subsystems, it can be seen that the organizational mechanism changes more frequently and the order degree shows a downward trend, which is the main reason why the IoT industry cannot develop in a coordinated manner. The development of industrial system synergy has contributed a lot.

4.3. Evaluation of the Collaborative Development Relationship of the Various Subsystems of the IoT Industry. On the basis of the overall collaborative measurement analysis of the IoT industrial system, a dynamic correlation analysis model of the synergy of the various subsystems within the IoT industrial system is constructed according to the gray system theory, namely, the GM (1, N) model.

4.3.1. Establishment of the Basic Data Sequence of Gray Evaluation of the Industrial System of the IoT. Based on the index data of the four subsystems of the IoT industrial system, this paper establishes the basic index data column of the gray system dynamic evaluation model. Based on the calculation of the index weights above, the dimensionless development sequence values of the four subsystems are calculated. This is the basic data column of the model. According to the basic data sequence method of establishing the gray dynamic evaluation model, the basic data sequence of the development of each subsystem of the IoT industry is obtained, as shown in Figure 9.

4.3.2. Analysis of the Development Capability of Each Subsystem of the IoT Industry. The coefficients a11, a22, a33, and a44 are the self-development coefficients of the integration, technological innovation, structural adjustment, and organizational mechanism of the IoT industrial system. Among them, the self-development coefficient of the fusion subsystem $a_{11} > 0$ indicates that the fusion subsystem has the ability of self-development. The main reason is that the basic data of the fusion subsystem is showing a gentle

upward trend, which also verifies the upward trend of the orderliness of the fusion subsystem. The self-development coefficient of the technological innovation subsystem $a_{22} > 0$ indicates that the subsystem has a good self-development ability, which verifies the rapid development trend of the technological innovation subsystem's orderliness, and the development level of the system's orderliness is higher than the other three subsystems. The order of science and technology expenditure and the order parameters of intangible assets has shown a rapid upward trend, mainly due to the continuous development of technological innovation in the IoT. The self-development coefficient of the structural adjustment mechanism subsystem $a_{33} > 0$ indicates that the structural adjustment mechanism subsystem has the ability to develop itself, but this ability is relatively weak. The organization mechanism subsystem $a_{44} > 0$ indicates that the organization mechanism subsystem currently has the ability to develop itself, and the order of system development is gradually improving. The gray model coefficients of each subsystem of the IoT industry are shown in Figure 10.

5. Conclusion

After establishing various mechanism models and solving them, this paper puts forward the conclusions of various mechanisms and application countermeasures based on the results of the solution and specially introduces a coordination manager who asks for the remaining system coordination gains to be responsible for the implementation of this series of coordination mechanisms. The system coordination manager should allocate the synergy gain to the members scientifically and induce the members to help each other through the incentive mechanism. At the same time, behavior supervision and information sharing supervision mechanisms are also needed. If the lack of supervision mechanism leads to poor service quality, customers will not choose the service provider of the network system but choose other IoT organization model enterprise services, and the operational stability of the IoT organization model network system will also be affected. After the orderliness of the structural adjustment mechanism reached its lowest point, it began to rise gradually. In the structural adjustment mechanism, the labor structure has the largest impact, followed by the technical structure. Then increasing the introduction of talents and expanding intellectual support will play a role in improving the orderliness of the structural adjustment mechanism. The orderliness of the organizational mechanism shows a downward trend after reaching the peak, indicating that the organizational mechanism that can lead the development of the IoT industry is not yet perfect. Improving the performance of industrial organizations and improving profitability have a greater effect on the orderliness of organizational mechanisms. We analyze the development capabilities of the subsystems themselves and the relationship between them through the gray dynamic correlation analysis model. At present, each subsystem has the self-development capability, and the fusion subsystem and the technological innovation subsystem have a unidirectional weak synergy relationship.

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during the current study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] J. Liu, H. Nishiyama, N. Kato, and J. Guo, "On the outage probability of device-to-device-communication-enabled multichannel cellular networks: an RSS-threshold-based perspective," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 1, pp. 163–175, 2016.
- [2] V. K. Chawla, A. K. Chanda, S. Angra, and S. Rani, "Effect of nature-inspired algorithms and hybrid dispatching rules on the performance of automatic guided vehicles in the flexible manufacturing system," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 41, no. 10, p. 391, 2019.
- [3] J. Myers, A. Savanth, R. Gaddh, D. Howard, P. Prabhat, and D. Flynn, "A subthreshold ARM cortex-M0+ subsystem in 65 nm CMOS for WSN applications with 14 power domains, 10T SRAM, and integrated voltage regulator," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 1, pp. 31–44, 2016.
- [4] A. Lamperski and J. C. Doyle, "The \mathcal{H}_2 control problem for quadratically invariant systems with delays," *IEEE Transactions on Automatic Control*, vol. 60, no. 7, pp. 1945–1950, 2015.
- [5] U. Orozco-Rosas, K. Picos, and O. Montiel, "Hybrid path planning algorithm based on membrane pseudo-bacterial potential field for autonomous mobile robots," *IEEE Access*, vol. 7, pp. 156787–156803, 2019.
- [6] Y. He, N. Zhao, and H. Yin, "Integrated networking, caching, and computing for connected vehicles: a deep reinforcement learning approach," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 1, pp. 44–55, 2018.
- [7] S. Pape and K. Rannenber, "Applying privacy patterns to the Internet of things' (IoT) architecture," *Mobile Networks and Applications*, vol. 24, no. 3, pp. 925–933, 2019.
- [8] H. Reyserhove and W. Dehaene, "A differential transmission gate design for minimum energy sub-10-pJ/cycle ARM cortex-M0 MCUs," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 7, pp. 1904–1914, 2017.
- [9] S. Sabau and N. C. Martins, "Youla-like parametrizations subject to QI subspace constraints," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1411–1422, 2014.
- [10] Y. Lu, C. Cheng, J. Yang, and G. Gui, "Improved hybrid precoding scheme for mmWave large-scale MIMO systems," *IEEE Access*, vol. 7, pp. 12027–12034, 2019.
- [11] Y. Zhao, X. Liu, G. Wang, S. Wu, and S. Han, "Dynamic resource reservation based collision and deadlock prevention for multi-AGVs," *IEEE Access*, vol. 8, pp. 82120–82130, 2020.
- [12] N. Zhang, P. Yang, J. Ren, D. Chen, L. Yu, and X. Shen, "Synergy of big data and 5G wireless networks: opportunities, approaches, and challenges," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 12–18, 2018.
- [13] V. Casola, A. De Benedictis, M. Erascu, J. Modic, and M. Rak, "Automatically enforcing security SLAs in the cloud," *IEEE Transactions on Services Computing*, vol. 10, no. 5, pp. 741–755, 2017.

- [14] S. Paul, V. Honkote, R. G. Kim et al., "A sub-cm³ energy-harvesting stacked wireless sensor node featuring a near-threshold voltage IA-32 Microcontroller in 14-nm tri-gate CMOS for always-ON always-sensing applications," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 4, pp. 961–971, 2017.
- [15] F. Lin, M. Fardad, and M. R. Jovanovic, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2426–2431, 2013.
- [16] A. Asef-Vaziri and M. Kazemi, "Covering and connectivity constraints in loop-based formulation of material flow network design in facility layout," *European Journal of Operational Research*, vol. 264, no. 3, pp. 1033–1044, 2018.
- [17] J. Ni, X. Lin, and X. S. Shen, "Efficient and secure service-oriented authentication supporting network slicing for 5G-enabled IoT," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 3, pp. 644–657, 2018.
- [18] F. Conti, R. Schilling, P. D. Schiavone et al., "An IoT endpoint system-on-chip for secure and energy-efficient near-sensor analytics," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 9, pp. 2481–2494, 2017.
- [19] K. Dvijotham, E. Todorov, and M. Fazel, "Convex structured controller design in finite horizon," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 1–10, 2015.
- [20] W. Xing, L. Peihuang, Y. Jun, Q. Xiaoming, and T. Dunbing, "Intersection recognition and guide-path selection for a vision-based AGV in a bidirectional flow network," *International Journal of Advanced Robotic Systems*, vol. 11, no. 3, p. 39, 2014.
- [21] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8440–8450, 2018.
- [22] R. Ranchal, B. Bhargava, P. Angin, and L. B. Othmane, "Epics: a framework for enforcing security policies in composite web services," *IEEE Transactions on Services Computing*, vol. 12, no. 3, pp. 415–428, 2019.

Research Article

Stepped Frequency Multiresolution Digital Signal Processing

Qunying Chen 

Xi'an Peihua University, Xi'an, Shaanxi 710125, China

Correspondence should be addressed to Qunying Chen; 150444@peihua.edu.cn

Received 29 April 2021; Revised 26 May 2021; Accepted 31 May 2021; Published 9 June 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Qunying Chen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of radar industry technology, the corresponding signal processing technology becomes more and more complex. For the radar with short-range detection function, its corresponding signal mostly presents the characteristics of wide bandwidth and multiresolution. In the traditional data processing process, a large number of signals will interfere with the signal, which makes the final signal processing difficult or even impossible. Based on this problem, this paper proposes a principal component linear prediction processing algorithm based on clutter suppression processing on the basis of traditional signal processing algorithm. According to the curve characteristics of the data returned by the target detected by the signal, through certain image signal measurement and transformation, the clutter can be effectively suppressed and the typical characteristics of the corresponding target curve can be enhanced. For the convergence problem of signal processing and the corresponding image chromatic aberration compensation problem, this paper will realize the chromatic aberration compensation of the corresponding target echo image based on the radial pointing transverse mode algorithm and enhance the convergence speed of the whole algorithm system. In the experimental part of this paper, the optimization algorithm proposed in this paper is compared with the traditional algorithm. The experimental results show that the algorithm proposed in this paper has obvious advantages in the convergence of signal processing and antijamming performance and has the promotion value.

1. Introduction

With the rapid development of information technology and electronic technology, radar technology has been further developed. The range resolution of common low altitude detection radar mainly depends on the bandwidth of the transmitted signal, the corresponding broadband data signal processing technology and antijamming technology [1–3]. At present, the traditional stepped frequency multiresolution digital signal processing technology in low altitude radar detection technology mainly focuses on the theoretical level, and the efficient and reasonable digital signal processing technology is still the problem of low altitude radar signal processing [4, 5]. At present, the factors that affect the signal echo of low altitude detection radar include the direct coupling of signals between antenna transceivers of low altitude detection radar, the large reflection on the ground and underground of the signal generating place, and a variety of irresistible reflection sources near the transmitting place, and the interference

transmitting source mainly comes from the complex and huge DSP printed board. The corresponding data signal needs to be converted between different frequency domains when processing the corresponding data signal, which will cause the sidelobe blocking phenomenon between the corresponding signals, thus forming a strong interference [6–8].

The traditional stepped frequency multiresolution digital signal processing mainly focuses on the system hardware design optimization and echo signal optimization processing. In the corresponding hardware level, it is not the key problem to be considered in this paper. For the digital signal processing level, it mainly deals with the close range digital clutter. The corresponding traditional processing methods include average method digital signal processing algorithm, digital signal principal component analysis algorithm, and corresponding digital signal linear prediction algorithm [9–11]. In the corresponding digital signal average processing algorithm, it is limited by the detection environment, which requires the corresponding detection environment. If

it is the ground, it requires the ground to be as smooth as possible. When the detection environment is low altitude, it requires less interference. After meeting the relevant environmental requirements, it averages the corresponding elements of the echo, so as to filter a large number of clutters. In essence, it is to arrange and combine a large number of original data returned by echo in a certain logical order and form a two-dimensional sequence, corresponding to each element in the sequence minus the average value one by one [12]. The corresponding digital signal principal component analysis algorithm assumes that the clutter presents a Gaussian distribution during detection. Under this assumption, the signal is decomposed by wavelet, so as to reduce the clutter component of radar detection signal. For the principal component of digital signal, it is decomposed by two-dimensional determinant, and finally the related clutter is suppressed [13]. For the digital signal linear prediction processing algorithm, it mainly uses the autoregressive average algorithm to process. In the process of processing, it takes the related signal as the reference signal and predicts the expected output of other signals. Under the corresponding mathematical model, it transmits the corresponding reference signal to the input of the digital signal processing algorithm. In this way, we can maximize the correlation between the corresponding algorithm processing ability and the corresponding processing effect [14]. Based on the above situation, it can be seen that the traditional digital signal processing technology has some problems, especially in the convergence and anti-interference degree of the algorithm which can not achieve a certain compromise. Based on this, this paper proposes a principal component linear prediction processing algorithm based on clutter suppression processing according to the curve characteristics of the data returned by the target detected by the signal, the clutter can be effectively suppressed, and the typical characteristics of the corresponding target curve can be enhanced through certain image signal measurement and transformation. For the convergence problem of signal processing and the corresponding image chromatic aberration compensation problem, this paper will realize the chromatic aberration compensation of the corresponding target echo image based on the radial pointing transverse mode algorithm and enhance the convergence speed of the whole algorithm system. In the experimental part of this paper, the optimization algorithm proposed in this paper is compared with the traditional algorithm. The experimental results show that the algorithm proposed in this paper has obvious advantages in the convergence of signal processing and antijamming performance and has the promotion value.

The structure of this paper is as follows: in the second section, the current research status of stepped frequency multiresolution digital signal processing algorithm is analyzed and discussed. In the third section of this paper, we will focus on the analysis of stepped frequency multiresolution digital signal processing algorithm, mainly on the analysis of principal component linear prediction processing algorithm based on clutter suppression processing and radius pointing transverse mode algorithm. In the fourth section of this paper, we will make a comparative experiment on the algorithm. Finally, this paper is summarized.

2. Related Work Analysis: Research Status of Stepped Frequency Multiresolution Digital Signal Processing Algorithm

The main technical difficulty in the field of stepped frequency low altitude radar detection is digital signal processing algorithm. Stepped frequency radar mainly transmits a group of wideband pulse signals by carrier frequency hopping. The corresponding echo signal processing technology becomes the difficulty of digital signal processing [15, 16]. Through the research and analysis of a large number of scientific research institutions and military research institutes, there are three main frequency step multiresolution digital signal processing algorithms, which are corresponding to the mean value method, digital signal principal component analysis algorithm, and corresponding digital signal linear prediction algorithm. The emphasis of the three algorithms is different, and the corresponding data signal processing technology has its own advantages and disadvantages. In the corresponding level of digital signal processing, it mainly includes digital signal average method, which is the so-called two-dimensional sequence processing algorithm, digital signal prediction algorithm, and other mainstream digital signal processing algorithms. The details of the corresponding algorithms are as follows: in the algorithm of digital signal average method, it is limited by the detection environment. If the detection environment is the ground, the ground should be as flat as possible. When the environment to be detected is low, less interference is needed. After meeting the relevant environmental requirements, the corresponding elements of the echo are averaged to filter out a large number of clutters. In essence, a large number of raw data returned by the echo are arranged and combined in a certain logical order to form a two-dimensional sequence. Each element in the corresponding sequence is subtracted from the average one by one. The algorithm has strong antijamming performance, but it has more strict requirements on the experimental environment, and most detection environments are complex and changeable [17–19]. The corresponding algorithm of digital signal principal component analysis assumes that the clutter presents Gaussian distribution in the corresponding detection. Under this assumption, the signal is decomposed by wavelet to reduce the clutter component of radar detection signal. The main components of digital signal are decomposed by two-dimensional determinant to suppress the correlated clutter. The algorithm has simple requirements for the environment and strong anticlutter processing ability, but the corresponding calculation is relatively complex, leading to serious convergence problems [20, 21]. For the linear prediction processing algorithm of digital signal, the autoregressive average algorithm is mainly used. In the process, the correlation signal is used as the reference signal to predict the expected output of other signals. In the corresponding mathematical model, the corresponding reference signal is transmitted to the input of digital signal processing algorithm, so as to maximize the correlation between the corresponding processing capacity and the

corresponding processing effect. The algorithm still has the problem of convergence [22–24]. Therefore, on the basis of the above analysis, combined with the actual situation of low altitude detection processing, it is of great significance to optimize the frequency step multiresolution digital signal processing algorithm.

3. Analysis and Research on Optimization Algorithm of Stepped Frequency Multiresolution Digital Signal Processing

This section mainly analyzes and studies the stepped frequency multiresolution signal processing algorithm. The main core algorithms are the principal component linear prediction processing algorithm and the radius pointing transverse mode algorithm. The corresponding algorithm architecture is shown in Figure 1. The corresponding algorithm architecture can be seen from the figure, and the corresponding algorithm hardware module architecture is also shown in the corresponding architecture diagram. As shown in Figure 1, the core algorithm is mainly based on the traditional signal processing algorithm, and the principal component linear prediction processing algorithm based on clutter suppression processing is proposed. According to the curve characteristics of the data returned by the target detected by the signal, through certain image signal measurement and transformation, the clutter can be effectively suppressed, and the typical characteristics of the corresponding target curve can be enhanced. So it gets rid of the disadvantages of single mean algorithm, principal component algorithm, and linear prediction algorithm. For the convergence problem of signal processing and the corresponding image chromatic aberration compensation problem, this algorithm is mainly based on the radial pointing transverse mode algorithm to realize the chromatic aberration compensation of the corresponding target echo image and enhance the convergence speed of the whole algorithm system. In the corresponding hardware design level, this paper mainly discusses the selection and configuration circuit design of FPGA and DSP and discusses the related design of auxiliary circuit.

In the corresponding experimental level, this paper first carries out the accurate modeling of the echo data signal and then carries out the experiment based on the correlation modeling. In the experimental part, this paper not only carries out the comparative experiments of four different algorithms for the processing of stepped frequency multiresolution digital signal, but also simulates the convergence and convergence speed of the four different algorithms based on MATLAB and the corresponding modeling. The simulation and experimental frameworks are shown in the figure.

3.1. Principal Component Linear Prediction Processing Algorithm Based on Clutter Suppression Processing. At the clutter suppression level, this section is mainly based on the linear prediction of principal components for clutter suppression. At the level of principal component processing, it mainly

analyzes and processes the collected echo signal from the perspective of two-dimensional determinant. At the level of principal component, the corresponding core processing ideas are as follows: based on the perspective of two-dimensional determinant analysis, it constructs and calculates the corresponding unrelated secondary fields, so as to analyze the relationship between the weights of the secondary fields. Taking the echo data as an example, the principal component analysis is carried out. Assuming that the corresponding echo data is *Data1*, the corresponding data is deconstructed based on the principal component. Formula (1) is the corresponding deconstruction formula. In the deconstruction formula, the corresponding mathematical symbol *n* represents the label of the corresponding data sampling point, the mathematical symbol *m* represents the number of data decomposition channels, and the corresponding *A_i* represents the *i*th target vector.

$$Data_n = A_{i-m} a_{i-m}^T = A_1 a_1 + A_2 a_2 + A_3 a_3 + \dots A_m a_m, \quad (1)$$

Based on formula (1), the determinant of echo data is processed with covariance. The corresponding *q* represents the projection size of the corresponding determinant in the corresponding direction after the corresponding echo data conversion, which represents the fluctuation degree of the echo data determinant in the corresponding direction and the corresponding energy size.

$$\partial_i = \frac{s_i^T s_i}{m}, \quad i = 1, 2, 3 \dots m. \quad (2)$$

Based on this, the corresponding principal component processing algorithm mode is determined as shown in Figure 2. From the figure, we can see that the corresponding processing details are as follows.

Step 1: take the corresponding echo data as the corresponding processing object, and the corresponding processing matrix calculation formula is shown in formula (3), where the corresponding *n* represents the corresponding echo data sampling points, and the corresponding *m* represents the corresponding echo data number.

$$Z_{i-m} = \begin{bmatrix} s_1 \\ s_2 \\ \dots \\ s_n \end{bmatrix}^T = \begin{bmatrix} z_{11} & \dots & \dots & z_{1m} \\ z_{21} & \dots & \dots & z_{2m} \\ \dots & \dots & \dots & \dots \\ z_{m1} & z_{m2} & \dots & z_{mm} \end{bmatrix}, \quad (3)$$

Step 2: use the traditional mean algorithm to process the echo data, and preprocess the corresponding echo data.

Step 3: based on the corresponding two-dimensional processing matrix, the corresponding echo data is processed by covariance matrix.

Step 4: analyze the corresponding eigenvalues and eigenvector values of the matrix after processing the corresponding echo data.

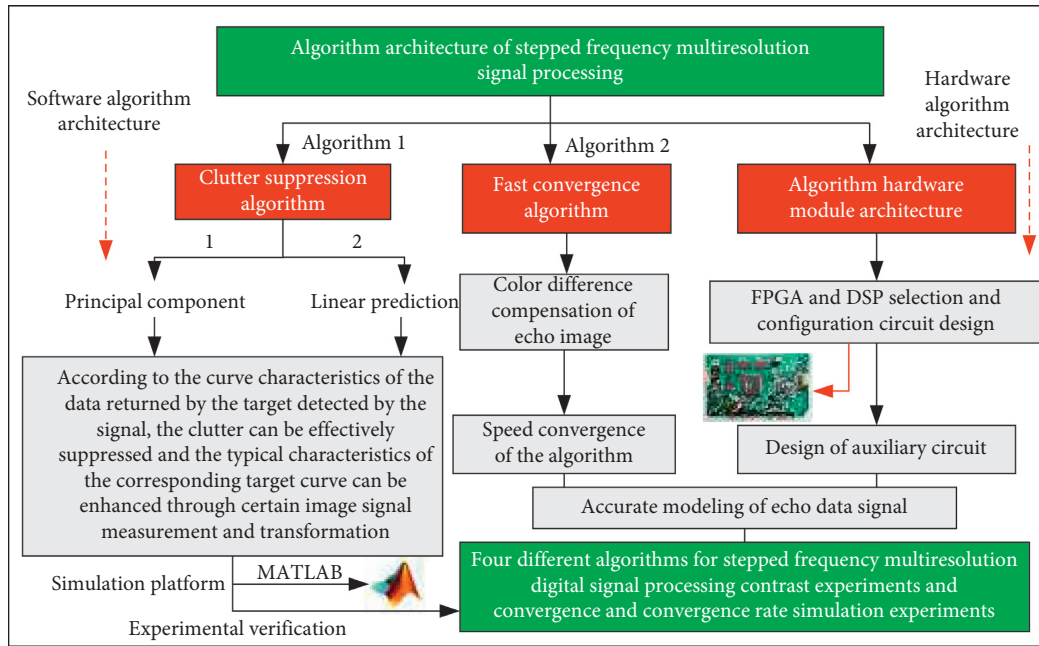


FIGURE 1: Architecture of stepped frequency multiresolution signal processing algorithm.

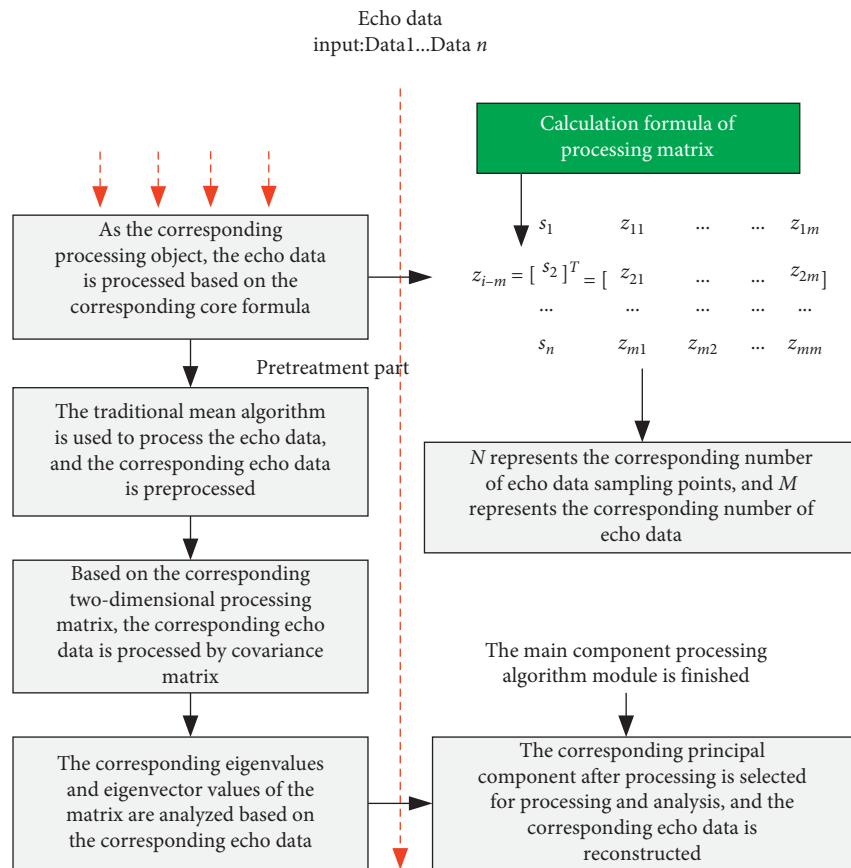


FIGURE 2: Algorithm operation flowchart of principal component processing algorithm module.

Step 5: select the corresponding principal component for processing and analysis, and reconstruct the corresponding echo data.

Based on the corresponding eigenvalues and eigenvector values obtained by the above principal component algorithm as the reference data of the linear prediction algorithm, the

corresponding echo data is further processed and analyzed, which mainly realizes the prediction of the current measurement data. Based on this, the corresponding core processing formula is shown in formula (4).

$$Z(i) = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mm} \end{bmatrix} * [i \ i-1 \ \dots \ i-m]. \quad (4)$$

For the complex interference factors faced by close range detection, the mathematical model is defined based on the above formula and the corresponding close range environmental factors, and the corresponding mathematical model is shown in Figure 3.

The echo data processed based on the principal component is taken as the corresponding reference data, and the corresponding bilateral data of the echo data matrix is selected as the linear prediction data of the current radar scan, and the number of corresponding reference data is set to W . When the number of reference data is small, the corresponding reference range will narrow, and the corresponding prediction effect will be greatly reduced. When the number of reference data is large, the detection difference of the corresponding radar in the corresponding area will be large. In this way, the corresponding data will lose its representativeness. Therefore, when dealing with the number of reference data, this section filters based on the data eigenvalues after the principal component preprocessing, so as to solve the problem of the number of reference data. In this section, the actual selection is generally between 10 and 20. Based on the above principle analysis, the corresponding linear prediction algorithm flow based on principal component preprocessing is shown in Figure 4. From Figure 4, the corresponding processing steps are as follows.

Step 1: recombine the collected data after preprocessing the corresponding principal components, rename the corresponding data matrix, and reanalyze all the new matrices.

Step 2: select a reasonable subspace from the above analysis results and recombine the corresponding information data.

Step 3: conduct linear bilateral processing analysis on the data in step 2 above, and the number of corresponding linear bilateral processing data is set to 6 (12 in total).

Step 4: use the least square method to predict the above data, so as to finally obtain the measured value and the corresponding estimated value, and based on the measured value and the estimated value, reduce the difference.

In order to solve the coupling interference between the equipment antennas corresponding to the short-range detection system, eliminate the external interference factors of nonalgorithm factors, so as to further optimize the above algorithm; this section designs the auxiliary design of antenna coupling interference elimination. The coupling

between antennas mainly exists in the fixed and stable clutter between the transmitting antenna and the receiving antenna. Based on the characteristics of this clutter, this section uses the air acquisition as the actual reference and subtracts the actual collected data. The corresponding processing steps are as follows.

Step 1: set up the corresponding short-range detection equipment, debug the corresponding parameters, and keep the interval between the receiving and transmitting antennas constant.

Step 2: shoot the corresponding equipment antenna to a clean environment without clutter interference, and collect data for the corresponding clean environment.

Step 3: preprocess the collected data to remove the interference of the corresponding target source. The collected echo data is mainly the mutual coupling direct wave between the receiving and transmitting antennas.

Step 4: average and store the corresponding data. The current reference data should be subtracted from the subsequent collected echo data so that the cleaner data can be obtained in the actual data processing of this algorithm.

3.2. Radius Pointing Transverse Mode Algorithm. In order to improve the convergence speed of the whole linear prediction algorithm and solve the unipolarity problem of the corresponding image problem, this section proposes the radius pointing transverse mode algorithm based on independent cell analysis algorithm. When the corresponding echo data is preprocessed by principal component, part of the signal is preprocessed at the same time. The corresponding transverse mode algorithm is used to iterate the corresponding data. The corresponding iteration termination condition is set as the absolute value threshold. Based on this, the inverse matrix of the corresponding echo data can be obtained. Based on the inverse matrix obtained above, the radius pointing transverse mode algorithm is used for depolarization multiplexing, so that the processing of echo data can get fast convergence. Based on this, the corresponding algorithm flow chart is shown in Figure 5. It can be seen from the figure that the corresponding algorithm flow details are as follows.

Step 1: preprocess and analyze the collected echo data.

Step 2: based on constant modulus algorithm, the preprocessed data are processed iteratively.

Step 3: repeat step 2 until the iteration condition is terminated (a certain threshold is met) to obtain the corresponding inverse matrix of echo data.

Step 4: based on the inverse matrix of echo data, the radius pointing transverse mode algorithm is used for depolarization multiplexing to achieve fast convergence.

3.3. Algorithm Hardware System Design. The corresponding hardware system is designed based on the above algorithm, and the corresponding hardware system architecture is shown in Figure 6. It can be seen from the figure that the

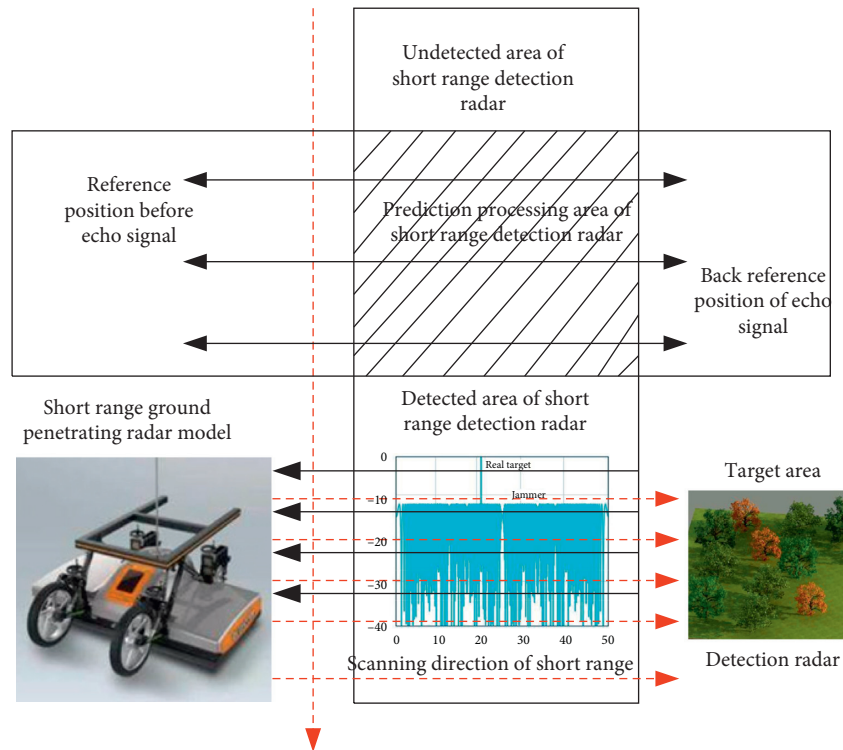


FIGURE 3: Mathematical model of linear prediction algorithm.

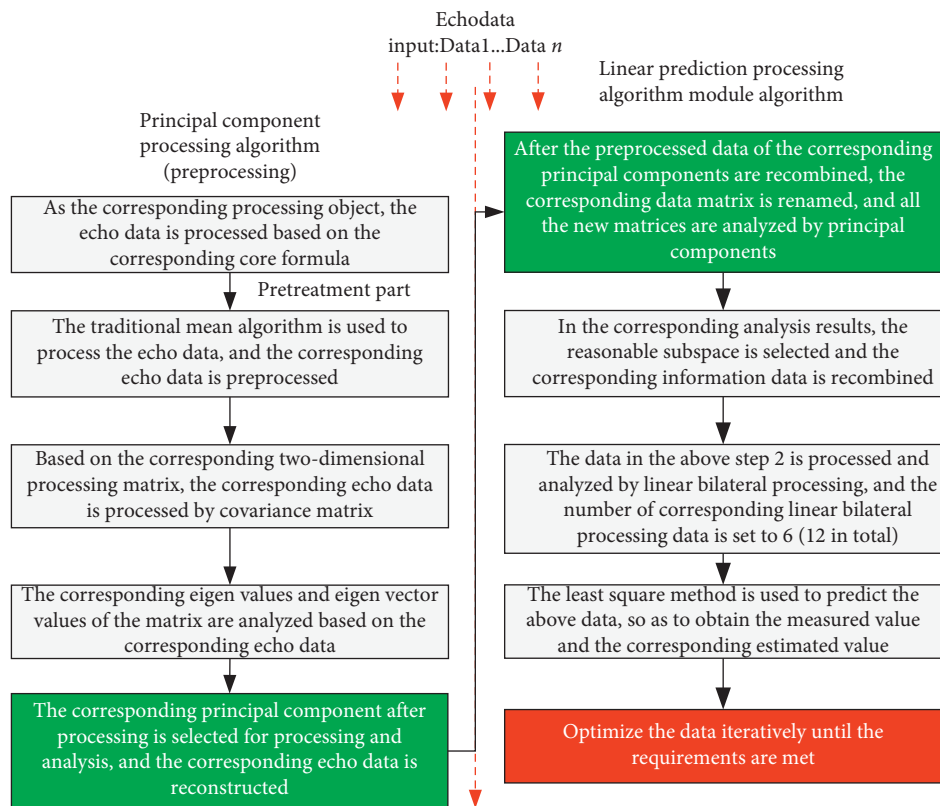


FIGURE 4: Algorithm operation flowchart of linear prediction processing algorithm module.

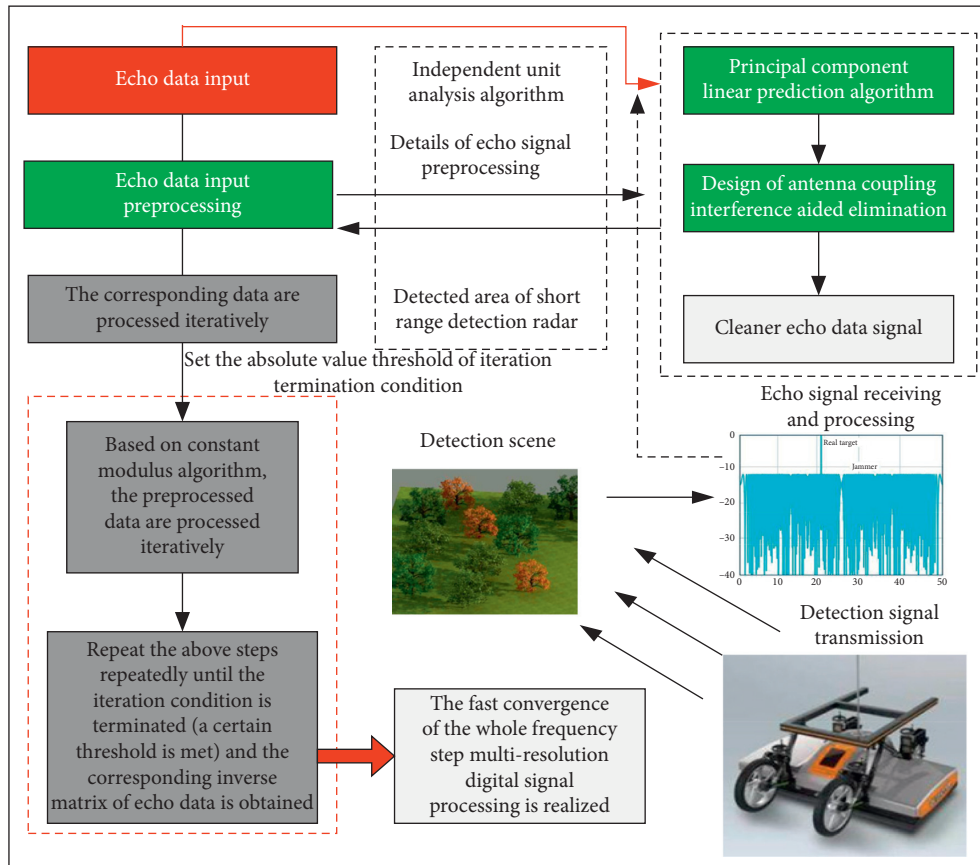


FIGURE 5: Flowchart of radius pointing transverse die algorithm.

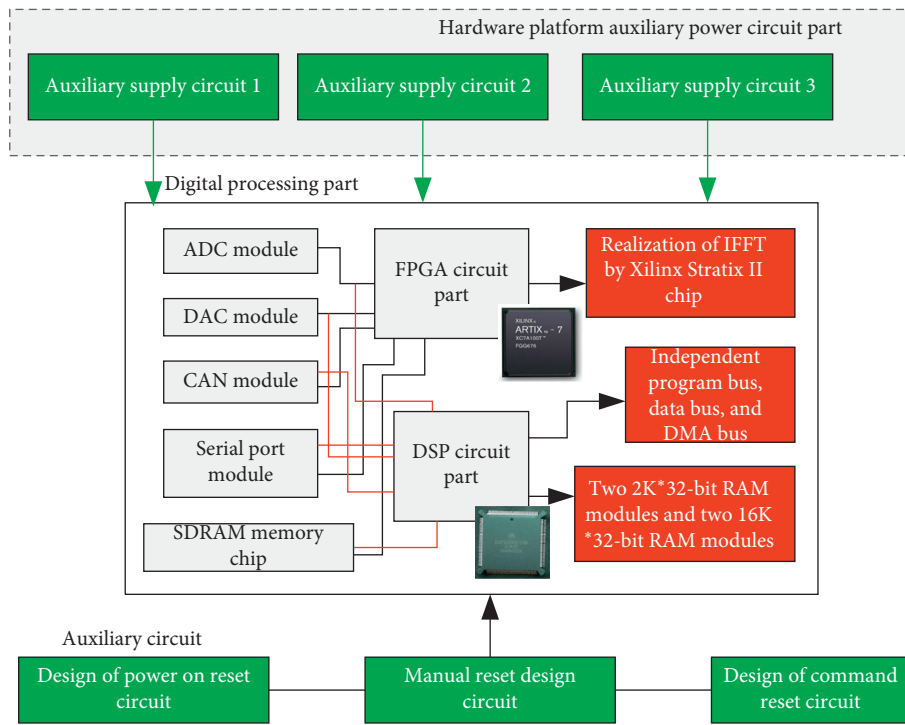


FIGURE 6: Hardware system architecture of stepped frequency multiresolution digital signal processing algorithm.

corresponding hardware modules include FPGA data processing module, DSP data processing module, and auxiliary data processing module. The corresponding data processing module mainly uses the Xilinx chip for data processing. At the same time, the whole data processing unit is configured with two SDRAM memory chips and two flash chips. The corresponding front-end signal acquisition module mainly uses the multichannel ADC module and the corresponding FPGA interface module.

In the corresponding data acquisition level, the relevant chips of ad company are mainly used, and the main models include ad9042, ad9631, ad8138, and ad4938. For the circuit composed of ad9631 and its related resistance and capacitance, its main function is DC coupling circuit, and its corresponding circuit needs to be equipped with emitter follower composed of operational amplifier. Its main function is to draw out the internal bias voltage corresponding to ad9042, and the adjustment of bias voltage is still mainly carried out through operational amplifier. In the ADC system, bias correction is needed. The corrected bias voltage and the input echo digital signal are synthesized by the broadband low noise operational amplifier and used as the analog input of ad9042. In the design of the above data acquisition circuit, the power supply part of the digital circuit should be separated from the analog part to prevent its impact on the ADC conversion speed.

The implementation of IFFT in the key FPGA circuit of the corresponding digital processing part mainly realizes FFT operation through the IP core designed by Altera company. In the aspect of selecting the corresponding DSP, this paper mainly selects the TMS320C3X series of Texas Instruments. Its corresponding structure is relatively simple, and the corresponding peripheral devices are relatively few. At the same time, it has the advantages of high-quality floating-point, low power consumption, rich registers, and so on. At the same time, the corresponding hardware resources of the processor are also very rich. It has two 2K*32-bit RAM modules and two 16K*32-bit RAM modules. Its independent program bus, data bus, and DMA bus enable the processor to realize parallel processing.

In the above hardware architecture system, the corresponding auxiliary circuit design mainly includes power management circuit module, clock circuit design part, and reset circuit design part. The design of the corresponding power management module mainly considers the power supply to the corresponding FPGA and the corresponding DSP processor. In the aspect of clock design, the corresponding DSP, FPGA, ADC, and all kinds of memory circuits in the hardware board are homologous with the external clock. Considering the self-test function of the whole system, crystal oscillator is set in the corresponding printed circuit board. The corresponding reset circuit design level mainly includes three types of reset design: power on reset design, manual reset design, and corresponding command reset design.

4. Signal Modeling and Experimental Analysis

4.1. Echo Signal Modeling. In order to better analyze the echo signal and prepare for the following experiments, this

section will model and process the echo signal. The theoretical analysis shows that the echo signal is mainly composed of three parts: external interference clutter, external interference noise, and the corresponding short-range target signal. When the corresponding equipment collects the echo data signal, it can be directly reflected in the signal spectrum. According to practical experience, the corresponding clutter signal can be attributed to the clutter interference directly coupled from the near air or ground underground. To a certain extent, the modeling of echo signal can be understood as the modeling of clutter signal, which corresponds to the clutter in the near air, underground, and ground. It is mainly caused by the debris in the near environment. The characteristic of this kind of clutter signal is that its change in the corresponding amplitude point is not very large. The corresponding clutter hidden in the ground or in the air is closely related to the target signal source. The corresponding clutter is often caused by the serious distortion of the underground or near space environment or the uneven distribution of the media in the corresponding cavity. The corresponding clutter is a rapidly changing process. Therefore, the model corresponding to this kind of clutter is random. The signals generated by the corresponding target source are usually signals with small amplitude. The corresponding other types of clutter are usually random clutter, which are generally set as the standard Gaussian white noise mathematical model, and are basically processed and analyzed by this model in the actual model processing. The corresponding coupling signals between antennas are pre-processed by subtraction algorithm. Based on the above analysis, the corresponding mathematical model of echo signal is established as shown in the following formula:

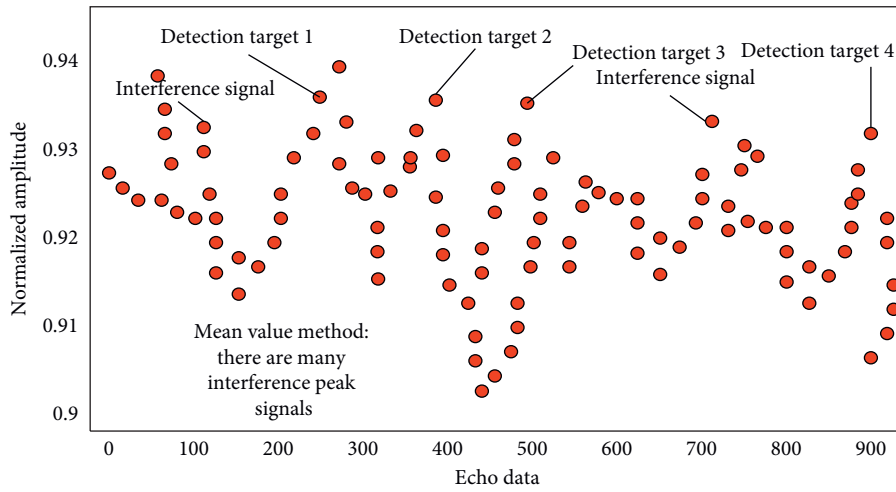
$$D = N_1 + N_2 + dN_3. \quad (5)$$

The experimental environment is set as follows: the experimental data is collected by a research institute, the corresponding experimental environment is a close ground, and there are holes and small stones of different specifications on the corresponding ground. The experiment is based on the same hardware system, and the data is collected along the test site in the way of horizontal movement, and the corresponding moving speed is set to uniform. In the actual comparative experiment, the experimental condition variables of the two algorithms are kept unchanged. In experiment scenario 1, the signal-to-noise ratio (SNR) is introduced to distinguish the experimental results. As shown in Table 1, the SNR values of different algorithms are shown. From the table, it can be seen that the algorithm proposed in this paper has obvious advantages over the single mean algorithm, principal component algorithm, and linear prediction algorithm. It has higher clutter suppression effect and higher SNR in scenario 1.

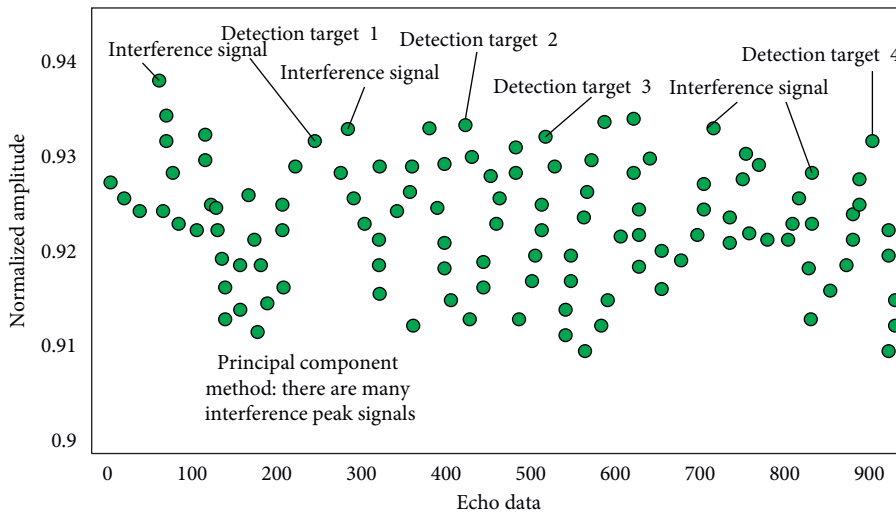
In experiment scenario 2, metal interferences are added on the basis of experiment scenario 1, and the corresponding radar hardware equipment continuously detects the detected target for five times in the test area. The amplitude accumulation diagrams of corresponding different algorithms are shown in Figures 7(a)–7(d). It can be seen from the figure

TABLE 1: Calculation results of SNR of different algorithms in scenario 1.

Digital signal processing algorithm	Signal-to-noise ratio
Raw data 1	-15.345
Mean method	8.987
Principal component algorithm	10.427
Linear prediction algorithm	15.293
An improved algorithm is proposed in this paper	30.114
Raw data 2	-20.351
Mean method	18.086
Principal component algorithm	22.135
Linear prediction algorithm	25.901
An improved algorithm is proposed in this paper	41.231
Raw data 3	-22.131
Mean method	20.861
Principal component algorithm	22.904
Linear prediction algorithm	25.619
An improved algorithm is proposed in this paper	33.891



(a)



(b)

FIGURE 7: Continued.

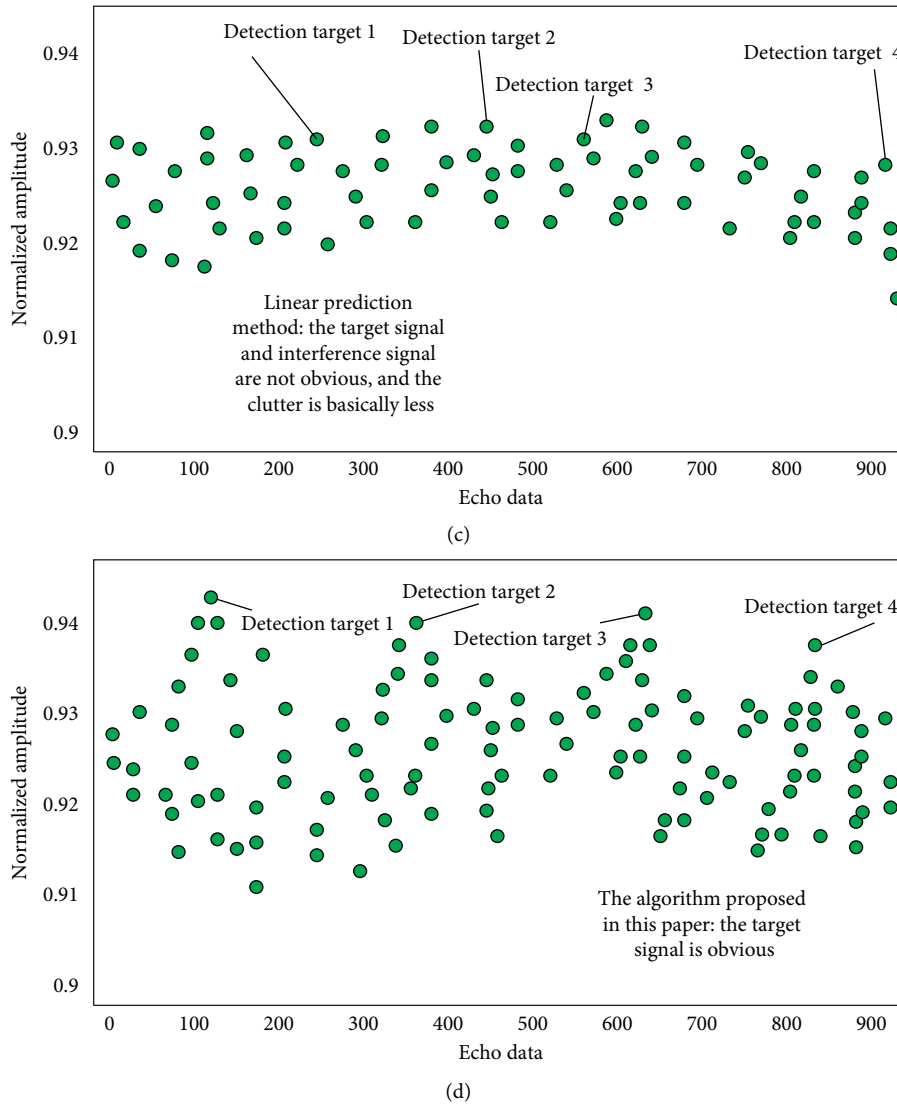


FIGURE 7: (a) Stacking chart of processing results under mean algorithm. (b) Stacking diagram of processing results under principal component algorithm. (c) Stacking chart of processing results under linear prediction algorithm. (d) The stack diagram of processing results under the algorithm proposed in this paper.

that compared with other algorithms, the algorithm proposed in this paper can obviously process five peak targets.

The SNR calculation table based on experimental scenario 2 is shown in Table 2. It can be seen from the table that the algorithm proposed in this paper has obvious advantages over the single mean algorithm, principal component algorithm, and linear prediction algorithm. It has higher clutter suppression effect and higher SNR in scenario 2.

In order to verify the convergence rate of the algorithm, the simulation is carried out based on the simulation framework shown in Figure 8 below, and the corresponding simulation tool is MATLAB. The simulation architecture shown in the figure is used to simulate and analyze the convergence of different algorithms.

In the corresponding simulation architecture diagram, the simulation is mainly divided into four channels, corresponding to mean value method, principal component

method, linear measurement method, and the algorithm proposed in this paper. In the actual simulation process, the principle of control variable method is followed, and the corresponding data source is guaranteed to be the same, and the corresponding data processing link is guaranteed to be the same. The only difference is the data processing module.

The convergence speed diagram of the four algorithms simulated based on the above simulation frame diagram is shown in Figure 9. From the figure, it can be seen that the convergence speed of the algorithm in this paper is faster than that of the other three algorithms. The algorithm proposed in this paper can reach the convergence point faster than the other three algorithms. At the same time, the curve trend in the graph can be seen. With the increase of data quantity, the convergence speed advantage of the proposed algorithm will be more obvious.

TABLE 2: SNR calculation results of different algorithms in scenario 2.

Digital signal processing algorithm	Signal-to-noise ratio
Raw data 1	-11.547
Mean method	4.111
Principal component algorithm	5.192
Linear prediction algorithm	8.361
An improved algorithm is proposed in this paper	20.324
Raw data 2	-15.339
Mean method	8.774
Principal component algorithm	15.934
Linear prediction algorithm	20.331
An improved algorithm is proposed in this paper	25.143
Raw data 3	-20.302
Mean method	11.213
Principal component algorithm	18.201
Linear prediction algorithm	21.354
An improved algorithm is proposed in this paper	30.119

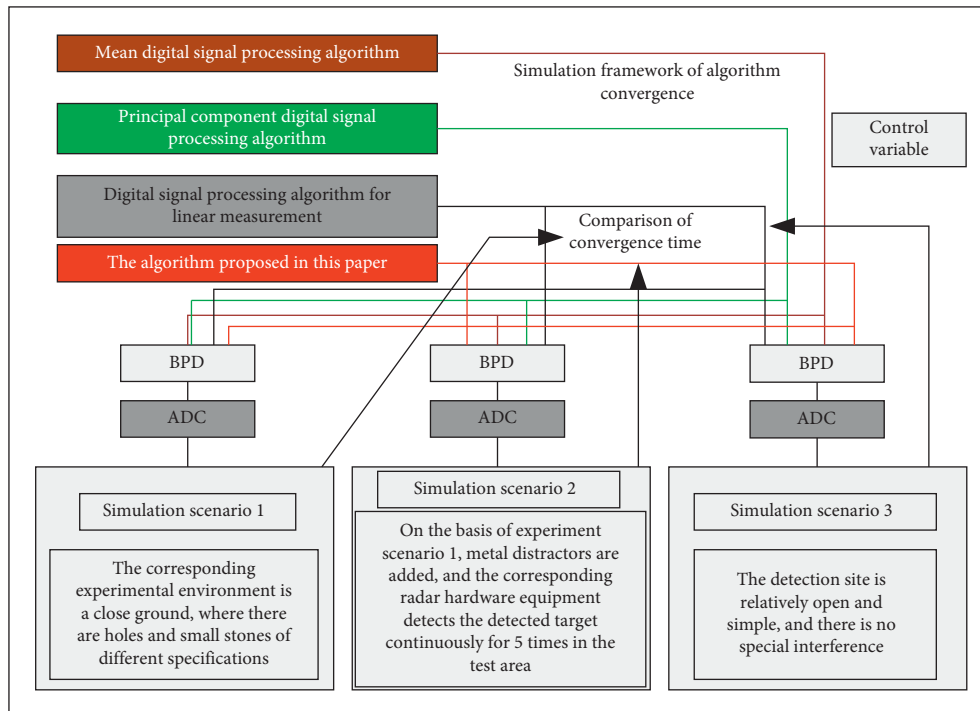


FIGURE 8: Simulation framework of algorithm convergence.

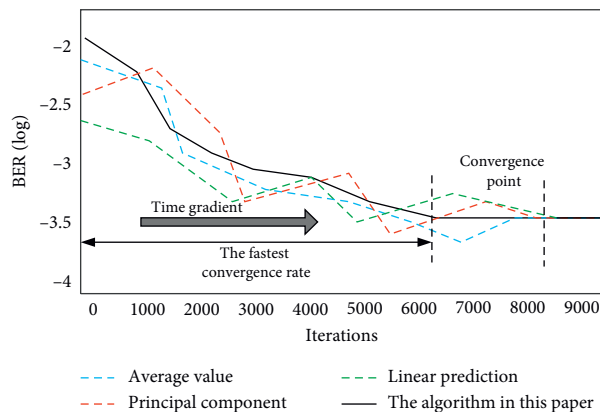


FIGURE 9: Curve of convergence rate of the algorithm.

Based on the data results of scenario 1 and scenario 2, we can draw the following conclusions: the single mean algorithm, principal component analysis, and linear prediction algorithm can filter most of the clutter when removing the clutter of the echo signal, but it still has the situation that the clutter can not be filtered in the face of complex environment, and it is subject to the relatively large correlation level. Compared with the single algorithm mentioned above, the algorithm proposed in this paper has obvious clutter filtering advantages, and its corresponding convergence speed is also very fast. Therefore, based on the above experimental results, the proposed algorithm has obvious processing advantages and convergence advantages.

5. Conclusions

This paper mainly analyzes the current situation and disadvantages of stepped frequency multiresolution digital signal processing technology. Based on the traditional signal processing algorithm, this paper proposes a principal component linear prediction processing algorithm based on clutter suppression processing. According to the curve characteristics of the data returned by the target detected by the signal, through certain image signal measurement and transformation, the clutter can be effectively suppressed and the typical characteristics of the corresponding target curve can be enhanced. For the convergence problem of signal processing and the corresponding image chromatic aberration compensation problem, this paper realizes the chromatic aberration compensation of the corresponding target echo image based on the radial pointing transverse mode algorithm and enhances the convergence speed of the whole algorithm system. In the experimental part, the proposed optimization algorithm is compared with the traditional algorithm. The experimental results show that the proposed algorithm has obvious advantages in the convergence of signal processing and antijamming performance and has the promotion value. This paper will focus on the optimization of convergence and antijamming for large-scale echo data.

Data Availability

No data were used to support the findings of the study.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The author acknowledges the scientific research projects of Xi'an Peihua University in 2020, Design of Wearable Motion Data Display System Based on Bluetooth (PHKT2029).

References

- [1] H. Iker and C. Zdemir, "Adaptation of stepped frequency continuous waveform to range-Doppler algorithm for SAR signal processing," *Digital Signal Processing*, vol. 106, no. 4, p. 102826, 2020.
- [2] C. L. Rino, K. M. Groves, C. S. Carrano et al., "Digital signal processing for ionospheric propagation diagnostics," *Radio Science*, vol. 50, no. 8, pp. 11–23, 2015.
- [3] S. Li, M. Xue, T. Qing, C. Yu, L. Wu, and S. Pan, "Ultrafast and ultrahigh-resolution optical vector analysis using linearly frequency-modulated waveform and dechirp processing," *Optics Letters*, vol. 44, no. 13, p. 3322, 2019.
- [4] S. J. Tingay, P. J. Hancock, R. B. Wayth et al., "A multi-resolution, multi-epoch low radio frequency survey of the Kepler K2 mission Campaign 1 field," *The Astronomical Journal*, vol. 152, no. 4, pp. 112–121, 2016.
- [5] O. Lobachev, C. Ulrich, B. S. Steiniger, V. Wilhelm, V. Stachniss, and M. Guthe, "Feature-based multi-resolution registration of immunostained serial sections," *Medical Image Analysis*, vol. 35, no. 4, pp. 288–302, 2017.
- [6] M. Chagnon, M. Morsy-Osman, D. Patel, V. Veerasubramanian, A. Samani, and D. Plant, "Digital signal processing for dual-polarization intensity and inter-polarization phase modulation formats using Stokes detection," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 188–195, 2016.
- [7] R. Kuldvee, M. Kaljurand, and H. C. Smit, "Improvement of signal-to-noise ratio of electropherograms and analysis reproducibility with digital signal processing and multiple injections," *Journal of Separation Science*, vol. 21, no. 3, pp. 169–174, 1998.
- [8] M. Agostini, M. Allardt, A. M. Bakalyarov et al., "Improvement of the energy resolution via an optimized digital signal processing in GERDA Phase I," *European Physical Journal C*, vol. 75, no. 6, p. 255, 2015.
- [9] S. Schaefer, M. Gregory, and W. Rosenkranz, "Coherent receiver design based on digital signal processing in optical high-speed intersatellite links with M -phase-shift keying," *Optical Engineering*, vol. 55, no. 11, p. 111614, 2016.
- [10] Y. Zhao, B. Hu, Z. A. He et al., "Generation and coherent detection of QPSK signal using a novel method of digital signal processing," *Modern Physics Letters B*, vol. 32, no. 4, p. 1850103, 2018.
- [11] J. Yu, L. Wang, P. Liao et al., "Method of digital signal process aided by control signal for burst-mode coherent receivers," *Optical Engineering*, vol. 54, no. 11, p. 116103, 2015.
- [12] W. Nascimento Lopes, F. Isaac Ferreira, F. Aparecido Alexandre et al., "Digital signal processing of acoustic emission signals using power spectral density and counts statistic applied to single-point dressing operation," *IET Science, Measurement & Technology*, vol. 11, no. 5, pp. 631–636, 2017.
- [13] M. Mazur, J. Schröder, A. Lorences-Riesgo, T. Yoshida, M. Karlsson, and P. A. Andrekson, "Overhead-optimization of pilot-based digital signal processing for flexible high spectral efficiency transmission," *Optics Express*, vol. 27, no. 17, p. 24654, 2019.
- [14] T. Yellamraju, A. J. Magana, and M. Boutin, "Investigating students' habits of mind in a course on digital signal processing," *IEEE Transactions on Education*, vol. 62, no. 4, pp. 312–324, 2019.
- [15] M. Altman, "Ardutouch: an Arduino-compatible synthesizer: digital signal processing squeezed into an easy-to-build kit-[Resources_Hands on]," *IEEE Spectrum*, vol. 55, no. 12, pp. 21–22, 2018.
- [16] K. Zhang, Q. Zhuge, H. Xin, and W. Hu, "Performance comparison of DML, EML and MZM in dispersion-unmanaged short reach transmissions with digital signal

- processing,” *Optics Express*, vol. 26, no. 26, pp. 34288–34304, 2018.
- [17] K. V. Ajay, A. Navneet, and I. Arif, “Comparison of digital signal processing, feedback and feedforward compensation technique for dual polarization 128-QAM radio over fiber link,” *Optik*, vol. 174, no. 5, pp. 68–76, 2018.
- [18] R. Asif, R. Ahmad, and R. Basir, “Hybrid opto-digital signal processing in 112 Gbit/s DP-16QAM and DP-QDB transmission for long-haul large- \mathcal{A}_{eff} pure-silica-core fiber links,” *Photonic Network Communications*, vol. 32, no. 1, pp. 1–8, 2016.
- [19] Z. Wang, J. Zhang, N. Verma et al., “Reducing quantization errors for inner-product operations in embedded digital signal processing systems [Tips&Tricks],” *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 141–147, 2016.
- [20] Y. Liu, H. Xiong, C. Dong, C. Zhao, Q. Zhou, and S. Li, “Real-time signal processing in field programmable gate array based digital gamma-ray spectrometer,” *Review of Scientific Instruments*, vol. 91, no. 10, p. 104707, 2020.
- [21] W. Wang, H. Li, P. Zhao et al., “Advanced digital signal processing for reach extension and performance enhancement of 112 gbps and beyond direct detected DML-based transmission,” *Journal of Lightwave Technology*, vol. 37, no. 1, pp. 163–169, 2018.
- [22] P. S. Lee, C. S. Lee, and J. H. Lee, “Development of FPGA-based digital signal processing system for radiation spectroscopy,” *Radiation Measurements*, vol. 48, no. 4, pp. 12–17, 2013.
- [23] Q. Zhuge, M. Morsy-Osman, M. Chagnon, X. Xu, M. Qiu, and D. V. Plant, “Terabit bandwidth-adaptive transmission using low-complexity format-transparent digital signal processing,” *Optics Express*, vol. 22, no. 3, p. 2278, 2014.
- [24] X. Liu, S. Chandrasekhar, and P. J. Winzer, “Digital signal processing techniques enabling multi-Tb/s superchannel transmission: an overview of recent advances in DSP-enabled superchannels,” *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 16–24, 2014.

Research Article

Research on the Key Technologies of Network Security-Oriented Situation Prediction

Yikun Zhu ¹ and Zhiling Du ²

¹*School of Mechanical Engineering & Automation, Wuhan Textile University, Wuhan, Hubei 430000, China*

²*School of Economics and Management, East China JiaoTong University, Jiangxi, Nanchang 330013, China*

Correspondence should be addressed to Zhiling Du; 2018048125300096@ecjtu.edu.cn

Received 22 February 2021; Revised 7 April 2021; Accepted 19 April 2021; Published 31 May 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Yikun Zhu and Zhiling Du. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In today's increasingly severe network security situation, network security situational awareness provides a more comprehensive and feasible new idea for the inadequacy of various single solutions and is currently a research hotspot in the field of network security. At present, there are still gaps or room for improvement in network security situational awareness in terms of model scheme improvement, comprehensive and integrated consideration, algorithm design optimization, etc. A lot of scientific research investments and results are still needed to improve the form of network security in a long and solid way. In this paper, we propose a network security posture assessment model based on time-varying evidence theory for the existing multisource information fusion technology that lacks consideration of the problem of threat occurrence support rate over time and make the threat information reflect the law of time change by introducing a time parameter in the basic probability assignment value. Thus, the existing hierarchical threat posture quantitative assessment technique is improved and a hierarchical multisource network security threat posture assessment model based on time-varying evidence theory is proposed. Finally, the superiority of the proposed model is verified through experiments.

1. Introduction

With the popularity and development of computer and network technology, various network systems have been deeply embedded in the daily production and life of society and individuals [1]. In recent years, network attacks, information leakage, and other ongoing security incidents have exposed the serious problems facing network security [2]. There are many factors that cause the deterioration of the form of network security, both internal and external causes [3]. The first and foremost of the internal causes is the construction level of the network system, which varies with the level of practitioners, the complexity of the system, and whether financial services are involved [4]. For example, in January, the user names and passwords of many network service companies, such as Tianya Community, were made public in clear text, leading to the leakage of a large number of user password habits and e-mail addresses [5–10]. Various

security loopholes, such as cross-site attacks, injection, hijacking, and other problems are rife in the network services, and even strong international companies are not immune [11]. In addition to the problems arising from the construction of the network system, deeper problems come from the system platform and some important service programs have been found to have major security vulnerabilities one after another. Use to the use of third-party operating systems or service programs (such as servers and a series of service software), system builders usually cannot and do not have the ability to discover all potential problems and can only rely on software providers or even open source organizations [12–16].

Not to mention that some open source organizations are slow to respond to security and slow to update. Even the system vulnerabilities released by Microsoft are threatened by “zero-day attacks.” Many attackers use the time interval between vulnerability releases and system updates to

successfully implement a large number of attacks. Various sources of threats make various network systems, whether corporate intranet or public network services, full of risks [17]. Globally, the Internet is frequently attacked, network paralysis occurs from time to time, and the security of network systems is constantly under serious threat; in China, the types and numbers of various network attacks continue to grow, especially the proliferation of Trojan horses is the most serious [18]. Network theft, network economic crimes, and large-scale network attacks have posed serious security threats to China's network infrastructure and important network systems, greatly restricting national economic development and even endangering social stability and national security. Network Security Situation Evaluation (NSSE) is a key technology in network security situational awareness, and the processed security situational elements data will be entered into the network security situational evaluation model for comprehensive evaluation. The evaluation needs to consider comprehensiveness, multiple granularity, and so on [19].

However, due to technical limitations, market expectations, funding, and other issues, most products can solve a single threat or in some cases have better results against certain attacks. But overall there is no unified management and scheduling mechanism and they are unable to cooperate with each other, unable to communicate effectively, and unable to track multiple data sources. This cannot cope with more complex attacks or long-term, potential, and gradual infiltration attacks that cannot be captured. Some interrelated attacks are lost in the tens of thousands of management logs of different products, and a large number of "false positives" have made administrators numb and ignore valuable information. Therefore, the industry has produced a more secure and effective global grasp of the theory of cyber security conditions, and tools are urgently needed. It is hoped that by integrating all aspects of safety information, the current qualitative and safety information can be extracted and calculated. It even quantifies the value of network security status to help administrators take timely remedial and preventive measures. The time decay function is used in the model to solve the dynamic problem of trust, and the stability and scalability of the model are enhanced through the trust prediction mechanism. Finally, the model is applied to the protocol, and the experimental simulation is carried out. The result verifies the safety and effectiveness of the model. The research in this area has also become a hot concern direction, and this paper has done a lot of research work based on this and achieved certain results.

2. Related Work

Here, we present some related work, which combines both external information of monitoring devices and network system environment information to form a situational value that can reflect the network system. It is a comprehensive technology with high requirements for both mathematical method application and network model. At present, the general assessment methods mainly include the following analysis methods: gray evaluation method, fuzzy analysis

method, hierarchical analysis method, Delphi expert method, and so on [20]. In another work, based on a large amount of past data information, researchers used the connectivity between them to predict the future security trends in a certain period of time [21]. Just as network attacks follow certain steps and patterns, the same situation exists for network security situations [22]. In addition, scientific methods are used to discover links and laws, accurately predict various network attacks, protect network system security management, achieve a reasonable level of security control before security incidents occur, and conduct targeted preprocessing [23]. Situation prediction techniques mainly include the gray theory prediction method, Bayesian inference-based prediction, autoregressive moving average model prediction, machine learning algorithm prediction, and neural network prediction.

In another case, researchers are mainly to facilitate network administrators to intuitively understand the operating status and development trend of the entire network and to give people a more intuitive feeling [24]. It uses computer graphics and image processing technology to convert static data into dynamic images. The graphics are displayed on the computer screen to realize the interactive connection between the maintenance personnel of the network system and the security data behind. For example, the NGSOC situational awareness platform released by domestic 360 security company and the Cloud Shield situational awareness released by AliCloud provide good examples of network security situational visualization [25].

A network security situational awareness framework is a macro representation of the whole picture of situational awareness. It expresses the functional tasks of each aspect of situational awareness in abstract and general semantics, which helps to show the flow and development of the whole security posture. A good situational awareness framework not only carries the whole process of technology pointing but also shows a clear outline for the next step of development. According to the situational awareness foundation model, a more comprehensive reference model for network security situational awareness currently combines different aspects and types of security systems such as network firewalls, intrusion detection systems, and security audit systems.

2.1. Network Security Posture Framework System. From the collection of information elements, data preprocessing can understand the current network security situation and, at the same time, analyze and predict the future trend of network security changes. The rational analysis of the security situation combines system equipment and resources to achieve effective control, thereby bringing an intuitive and comprehensive response to decision makers. A typical NSSA framework is shown in Figure 1.

In this framework, the NSSA implementation is composed of five levels (stages), starting with security information collection work, understanding events through subsequent processing steps, in which the system provides real-time control feedback, and finally visualizing and

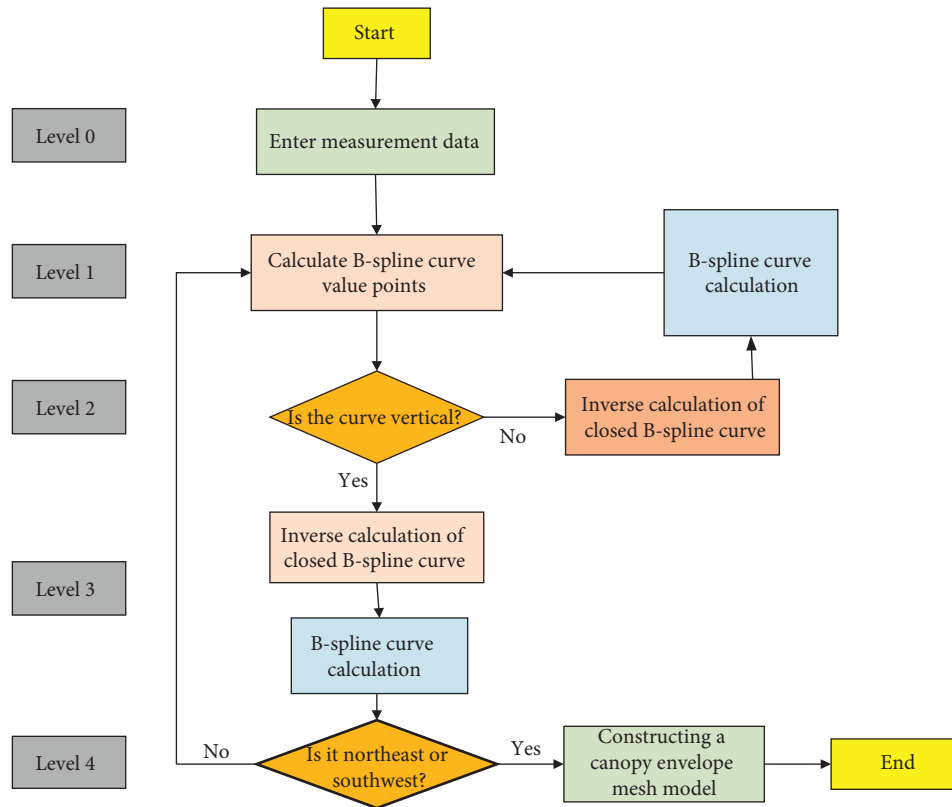


FIGURE 1: Typical NSSA framework.

analyzing the security posture based on human-computer interaction. Five of these processing levels are as follows:

Level 0 (data preprocessing): it includes a large amount of unstructured data, optional preprocessing levels for structured data and agile data, and obtains necessary data by obtaining contextual elements; it will become a part of nonuniform data preprocessing operations, such as data desensitization, cleaning, and impurity filtration.

Level 1 (event extraction): after obtaining the relevant element information, the information is extracted to form standardized events, as well as the rules and features of unified events.

Level 2 (situation assessment): correlation and data fusion operations are performed on the collected multisource data, and information events are understood using existing assessment algorithms to form a comprehensive security situational analysis report, which provides supplementary information for subsequent managers.

Level 3 (impact assessment): according to the previous period of the situation, analyze the possible future trends of change, combine the full range of resources and control of the network system, introduce expert experience into the decision, and give an evaluation strategy for the impact of future changes in the situation.

Level 4 (resource management, process control, and optimization): selecting objects for system monitoring,

real-time scheduling and allocation of occurring situation events, rational use of system resources, and maximizing efficiency.

2.2. Network Security Posture Indicator Construction Principles. Security posture indicator construction is an indispensable and fundamental part of the NSSA process, which is an indicative sign reflecting the security attributes of the perceived object and provides the basis for measurement and assessment for network security posture understanding and prediction. There are numerous indicators that can have an impact on the network posture, and screening out the posture indicators with typical data forms the data source for subsequent posture sessions and can provide reliable data support for the next step of assessment and prediction. The network security posture indicator system is a collection of indicators that can fully reflect the characteristics of network security, and the indicators are intrinsically linked to each other and play a complementary role, which is the basis for forming standardized objective quantitative analysis conclusions on network security evaluation, and it can reflect the basic appearance, quality, and level of network security of the perceived object. Therefore, the construction of network security posture indicators is of great significance for network security posture perception. The construction of the indicator system needs to refer to certain principles, specifically the construction of the network security posture indicator system from different dimensions.

- (1) Hierarchical classification principle: network security posture indicators are hierarchical (some are for the local network and some are for the large-scale network); these indicators have their different meanings in different environments, and the processing process varies, so they should be considered in a hierarchical classification.
- (2) Similarity principle: in a large macronetwork, there are many influencing factors to be considered, but there is no lack of similarity and cross-measurement data. For example, the distribution of data packets, the distribution size of data packets, and similar indicators should all be taken into consideration.
- (3) The principle of the combination of dynamic and static: because the index itself has its own characteristics, such as the distribution of network assets and equipment within a certain period of time, the network topology is not easy to change, while the network traffic information is always required to be processed, filtered, and collected.

Therefore, the indicators of these two categories should be treated differently and combined with the indicators' own characteristics for the corresponding combination. In a comprehensive consideration of network security, the nature of three aspects is usually used to describe the network security posture, specifically: network base operation, network vulnerability, and network threat level, as shown in Figure 2. When the computing device factor is in the locked state, it is determined that all secure data can be decrypted by any one or more applications located on the computing device. When the computing device is in the locked state, if it is determined in the determining step, there is at least some secure data. If it can be decrypted by at least one or more applications, the first indicator is displayed. These three aspects represent three dimensions, which basically cover all parts of the entity constituting the information network and can reflect the security posture of the network in a more comprehensive way, and many researchers and commercial organizations currently use this approach.

2.3. Network Security Prediction under Fuzzy Logic Theory.

There are a large number of fuzzy phenomena in nature, that is, there are things that are not so easy to make distinctions, for example, young and old, fat and thin, tall and short, and long and short; there will be a certain degree of fuzziness among these concepts. For these fuzzy concepts and phenomena, in recent years, with the development of the field of mathematics, gradually formed a modern applied mathematical science to solve fuzzy problems: fuzzy mathematics. This theory was proposed by Professor L. Zadeh of the University of California, a famous American scholar. Fuzzy logic is the key part of fuzzy mathematical theory, which uses the affiliation function instead of the classical Boolean truth-value logic, abandoning the traditional deterministic two-valued truth proposition in the tradition and generalizing this problem to affiliation, which will be more conducive to making innovations in uncertain and fuzzy problems. And

fuzzy logic has been widely used in related scientific fields and brings new directions for industrial development. A fuzzy set is defined as follows: for a general set U , any mapping $\mu_{\tilde{A}}$ from U to the interval $[0, 1]$ can determine a fuzzy subset of U , called a fuzzy set \tilde{A} on U , where the mapping $\mu_{\tilde{A}}$ is called the affiliation function of the fuzzy set \tilde{A} , and for an element x on U , $\mu_{\tilde{A}}(x)$ is called the affiliation of x to the fuzzy set \tilde{A} , which can also be written as $\tilde{A}(x)$.

$$\mu_A: U \Rightarrow [0, 1] \text{ for } x = \frac{1}{\mu_A}(x), \quad (1)$$

where U is called the theoretical domain of the fuzzy set A , μ_A is the affiliation function of this fuzzy set, any element x on U no longer has only two cases of belonging to A and not belonging to A , and each element x has an affiliation μ_A for $A(x)$. The degree of affiliation μ_A represents the degree, and a higher value of it indicates a higher degree of x belonging to A , while the opposite indicates a lower degree of x belonging to A . The fuzzy set is transformed into a normal set when the μ_A value domain takes the two endpoints of the closed interval $[0, 1]$.

If the theoretical domain $U = \{x_1, x_2, \dots, x_n\}$, then the fuzzy set \tilde{A} on U can be expressed as

$$\tilde{A} = \frac{\sum_{i=0}^n \mu_A(a_{ii})}{q_{ij} + p_{ij}}, \quad (2)$$

where $\mu_{\tilde{A}}(x_i)$ ($i = 1, 2, \dots, n$) is the affiliation degree and x_i is an element in the domain of the argument. When the affiliation degree is 0, the item can be omitted. When the theoretical domain U is a continuous set, then the fuzzy set \tilde{A} on U can be expressed as

$$\tilde{A} = \sum_{i=0}^n \prod_{j=1}^n \mu_A(x_{ij}). \quad (3)$$

It should be noted that summation and integration here are not in the original sense and are special representations of fuzzy sets. If the affiliation of all elements on the domain of the argument is given, it can be represented in a similar way to the classical mathematical set, also called here as the sequential pair representation.

$$\tilde{A} = \frac{\sqrt{\prod_q X_{II}}}{x_i}, \quad (4)$$

where each sequential pair consists of an element x and its corresponding affiliation $\mu_{\tilde{A}}(x)$ and the fuzzy set \tilde{A} contains all combinations thereof.

Since the fuzzy set and its affiliation function form a one-to-one correspondence, the operation of the fuzzy set is also inscribed and represented by the operation of the affiliation function.

Empty set: it is the set whose affiliation function is 0 for all elements x , denoted as \emptyset , i.e.,

$$\tilde{A}_{x_{II}} = \sum_{i=1}^q \chi_{ii} \Leftrightarrow \mu_A(x) = 0. \quad (5)$$

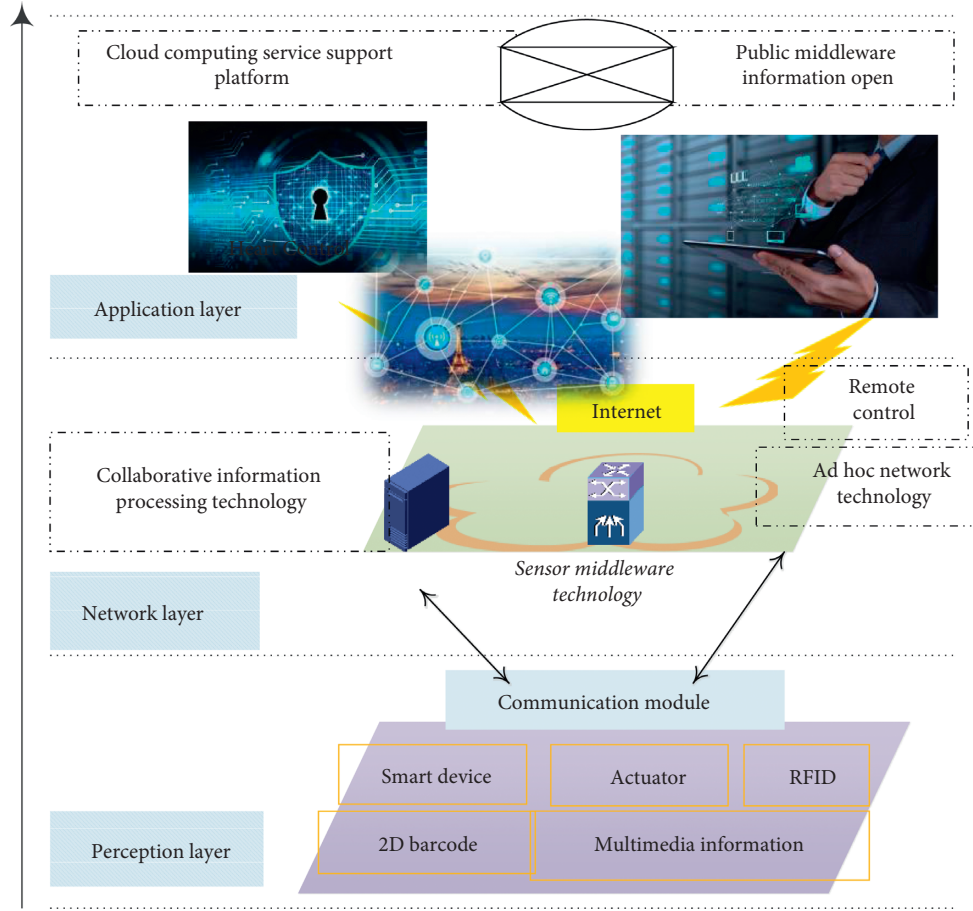


FIGURE 2: Network security influencing factors.

Equal sets: consider two fuzzy sets \tilde{A}, \tilde{B} ; if their affiliation functions are equal for all elements x , then \tilde{A}, \tilde{B} are also equal, i.e.,

$$\tilde{A} = \tilde{B} = \sum_{i=1}^q \mu_{ii}(x) \Leftrightarrow \mu_A(x) = 0. \quad (6)$$

Subsets: in the fuzzy sets \tilde{A}, \tilde{B} , \tilde{A} being a subset of \tilde{B} or \tilde{A} being contained in \tilde{B} means that for all elements x , there is $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$, denoted as $\tilde{A} \subseteq \tilde{B}$, i.e.,

$$\tilde{A} \in \tilde{B} \Leftrightarrow \mu_A(x) > 0. \quad (7)$$

Concatenation: for the concatenation set C of the fuzzy sets A, B , the affiliation function can be expressed as $\mu_{\tilde{C}}(x) = \max[\mu_A(x), \mu_B(x)]$, i.e.,

$$Q = \tilde{A} \cup \tilde{B} \Leftrightarrow \mu_A(x) = \max(\omega_A(x)) > 0. \quad (8)$$

Of course, the basic properties of fuzzy set operations (such as ordinary sets and fuzzy sets) also satisfy the power law, the exchange law, the union law, the absorption law, the distribution law, and the Morgan theorem. Ordinary set relations can only indicate that two factors are either related or unrelated, while fuzzy relations introduce uncertain quantities such as close relations, defined as follows: the direct product space $X \times Y = \{(x, y), x \in X, y \in Y\}$ of fuzzy

relations is a fuzzy set R on $X \times Y$, and the affiliation function R of $R(x, y)$ represents the degree of relationship between element x on X and element y on Y . The above fuzzy relation is also the simplest binary fuzzy relation, but of course it can be extended to form an n -element fuzzy relation. 50 counters are randomly distributed in an area of 1000×1000 , of which 10% are malicious binary. The converter uses a mobile model. The router's communication has a range of 250 m. The data source generator sends 4 cBR streams with a size of 512 B per second, the network bandwidth is 2 Mbps, and the simulation time is 300 s. If $R(x, y)$ takes the value of 0 or 1 special point, the fuzzy relation also degenerates to the ordinary set relation, as shown in Figure 3.

2.4. Neural Networks Predict Cybersecurity. Artificial neurons are the basic information processing units for ANN operation. Figure 4 represents the schematic diagram of the composition of ANN neuron, which has three main components: weighted adder, linear dynamic system, and nonlinear mapping function. x_i denotes the input from other neurons, $u(-)$ denotes the adjustable network connection weight, θ is the offset signal as the threshold, y_i denotes the output of the neuron, $(-)$ represents the basis function, which is a multiple-input single-output function, and $f(-)$ represents the activation function, which is a nonlinear mapping

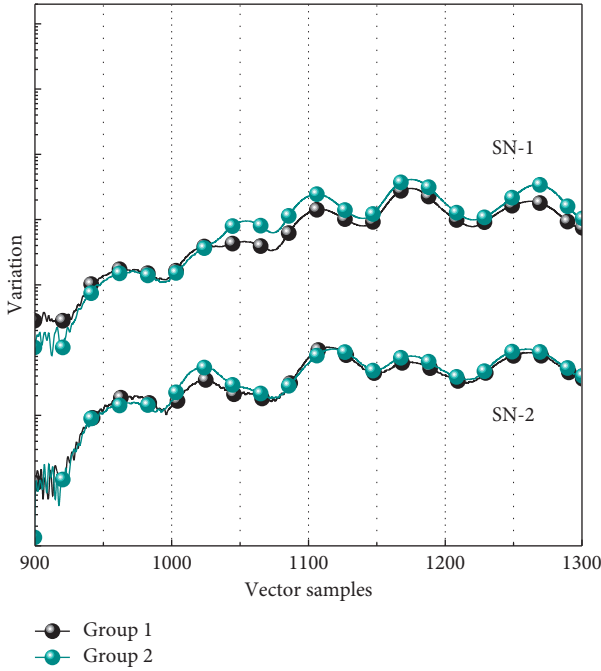


FIGURE 3: Fuzzy relationship transformation.

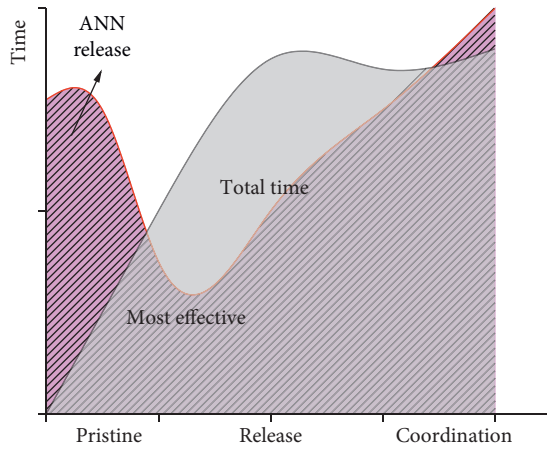


FIGURE 4: ANN general model.

of the output signal u of the basis function and is transformed into the specified range.

Sigmoid function: this function, also called S-shaped function, is most widely used in the ANN field. It is a strictly monotonic increasing function with smoothness and asymptotic characteristics. Its function expression is

$$f_{ij}(x) = \begin{cases} e^{(dQ_i/dt - dQ_j/dt)}, & \frac{dQ_i}{dt} < \frac{dQ_j}{dt}, \\ e^{(dQ^*/dt - dQ_j/dt)}, & \frac{dQ^*}{dt} < \frac{dQ_j}{dt}. \end{cases} \quad (9)$$

Among them, the parameter λ is the gain of the sigmoid function, which controls the slope of the function curve; generally, the larger the value of λ , the steeper the function

curve, and the function value domain is $(0, 1)$, which is often applied to the binary classification problem; however, this function also has the disadvantage that it is easy to oversaturate and produce the gradient disappearance phenomenon, which cannot complete the deep network training. Secondly, the mean value of the output of the function is not 0, which affects the operation efficiency of gradient descent.

Bipolar sigmoid function: the tanh function solves the problem of the existence of some unipolar S-shaped functions; it will transform the function value domain from $(0, 1)$ to $(-1, 1)$; compared with the original function, the convergence speed is faster, and the output average value is 0, as shown in Figure 5. Its function expression is

$$y = f(x) = \frac{\exp(\mu_A(x) + \mu_B(x))}{\exp(\mu_A(x) + \mu_B(x)) + \exp(\mu_A(x) - \mu_B(x))}. \quad (10)$$

Neural networks are composed of neurons as the basic unit. If a large number of neurons are connected in a topological structure according to certain rules to form a parallel distributed computing structure, the structure of neural networks as we know it in daily life is formed. In the following, several typical structures of neural networks are introduced according to the neural network connection pattern.

- (1) Single-layer perceptron: the single-layer perceptron is one of the simplest neural networks proposed by Frank Rosenblatt in 1957 and is mainly a structural model for binary classification problems. The function and mechanism of the single-layer perceptron is very simple and can be used to model simple logic functions with relatively few applications. The input layer as the perception layer in the figure has n neuron nodes, which are only responsible for the collection of external information but have no information processing capabilities. The input layer has n input signals, which constitute the input column vector X . The output layer, called the processing layer, is more specific in that each neural unit has information processing capability and the output constitutes the column vector Y . The connection weights between the input signals and the processing units are represented by the column vector W_j .
- (2) Feedforward network: feedforward neural network is the most widely used neural network structure model. It has the typical characteristics of the universal neural network structure; this network consists of three layers, which are the input layer, hidden layer, and output layer. In this network, the input signal is transmitted in one direction from the input layer to the output layer, the neurons in each layer are only connected to the neurons in the previous layer, there is no connection between the neurons in the same layer, and the whole network has no feedback, as shown in Figure 6.
- (3) Feedback-type network: feedback-type neural network establishes another connection between input

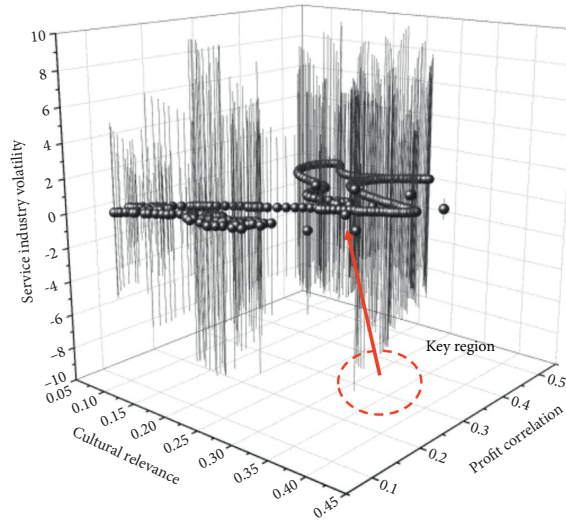


FIGURE 5: Matching of development speed between subsystems.

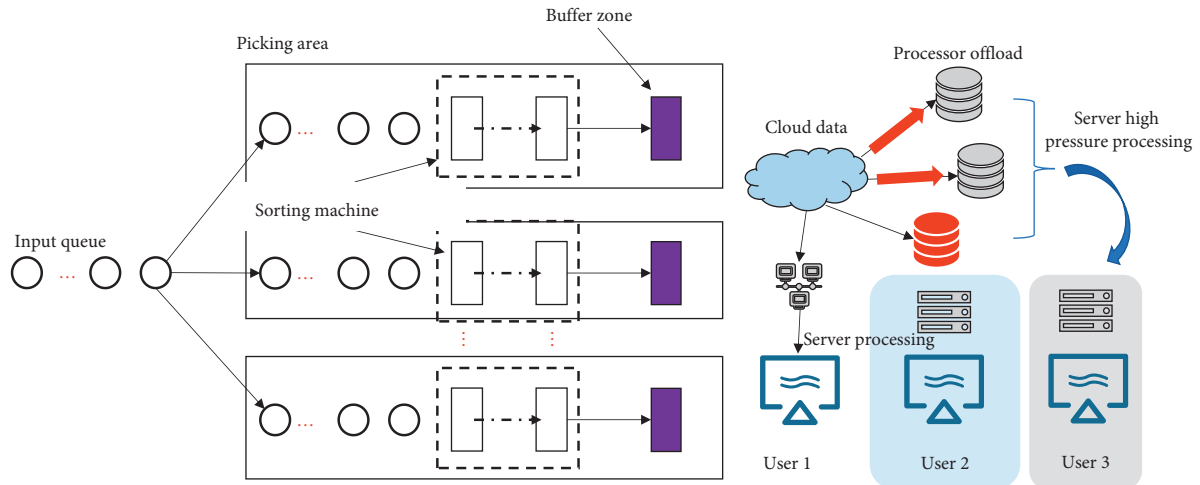


FIGURE 6: Neural network structure.

and output, that is, the signal of the output layer of the network is re-entered into the input layer through a feedback loop. Compared with the feed-forward neural network, this network has stronger computational ability and associative memory ability. This type of network can be used in real-time signal processing, system control, and other scenarios that require real-time adjustment according to the system state. Feedback-type neural networks can be further divided into full-feedback networks and partial-feedback networks.

The learning of neural networks also refers to the training process of the network structure parameters, through continuous learning to adjust and correct the parameters of the network in time to optimize its own performance and efficiency, so as to achieve the equilibrium conditions to adapt to the changes in the external environment. The learning methods of neural networks can be

divided into supervised learning and unsupervised learning. The training of neural networks mainly consists of forward propagation training and backward propagation training. Forward propagation training refers to the entire process from the input layer of the neural network, multiple hidden layers, and then the output layer to the output signal; the backward propagation mainly refers to the process of passing the error signal obtained from the output layer to the input layer in reverse order from back to front. Back-propagation mainly uses the gradient descent and error backpropagation (EBP) algorithms to adjust the weights of the network. In turn, the system error is obtained by comparing the output signal with the desired output signal, and the error of the latter layer is propagated forward layer by layer using chain derivatives to obtain the error signal of each layer, according to which the training network structure is adjusted, and the whole process of continuous cyclic training is the process of neural network learning. The training process can be terminated by artificially setting the

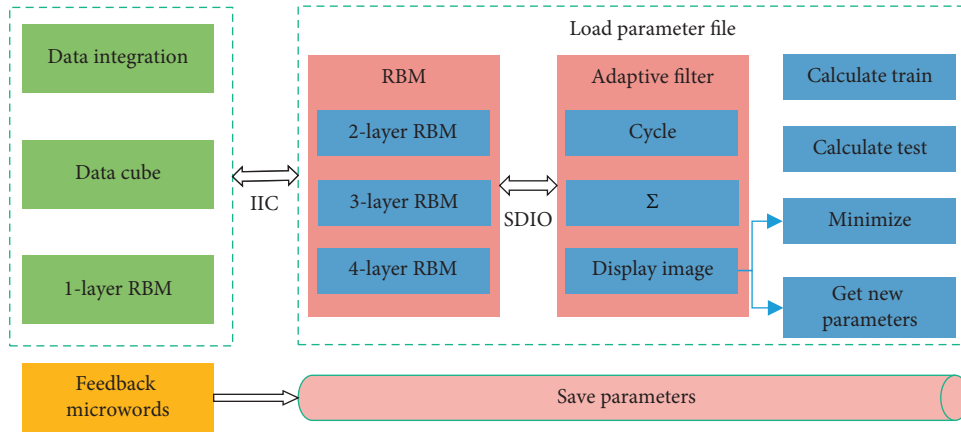


FIGURE 7: Neural network learning convergence.

termination conditions, either by setting the maximum number of iterations or by selecting the network convergence to a certain degree. If the number of iterations set is too small, as shown in Figure 7, it may cause the function to not converge well and the neural network to not fit well; if the number of iterations set is too large, it is easy to cause the function to overfit, resulting in insufficient generalization.

3. Analysis of Experimental Results

In order to verify the convergence efficiency of the PSO-LSTM prediction algorithm, the RNN prediction algorithm and LSTM prediction algorithm are combined here to train on the same test set and check their error convergence performance. The difference between the three algorithms is depicted in Figure 8. It can be found that the RNN has a large initial error, a large fluctuation in error convergence, and a long convergence time to the optimal error value during the training process, requiring more than 250 iterations. The LSTM algorithm with more hidden layers has a smaller initial MSE and faster convergence than RNN, requiring 150–200 iterations, but with more parameters, it is easy to fall into the problem of local optimum, and the error performance is only better than that of RNN overall. In contrast, the PSO-LSTM algorithm starts with a smaller error, and the subsequent convergence speed is significantly better than the previous two, with faster convergence time and higher stability. Therefore, it can be judged that introducing the particle swarm optimization algorithm into the LSTM neural network can indeed improve the convergence performance of the network.

In order to verify the prediction accuracy of the PSO-LSTM prediction algorithm, the three prediction methods are also analyzed and compared here, and Figure 9 shows the normalized posture prediction curves of the three for the next 50 days. The physical signal strength of the observation group decreases with time, while the control group is opposite and increases accordingly. Through the comparison, it is found that all three prediction algorithms have better prediction effect on the security posture in the future period, but PSO-LSTM can better track and analyze the security posture trend, overcoming the problem that the original

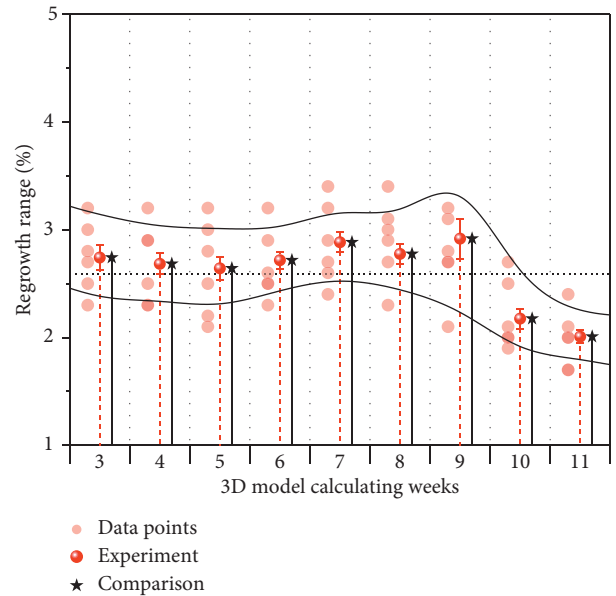


FIGURE 8: Comparison of MSE convergence speed by algorithm.

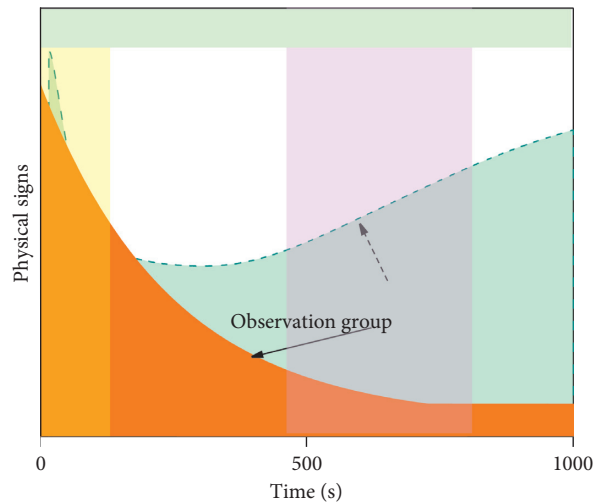


FIGURE 9: Security posture prediction curve for each algorithm.

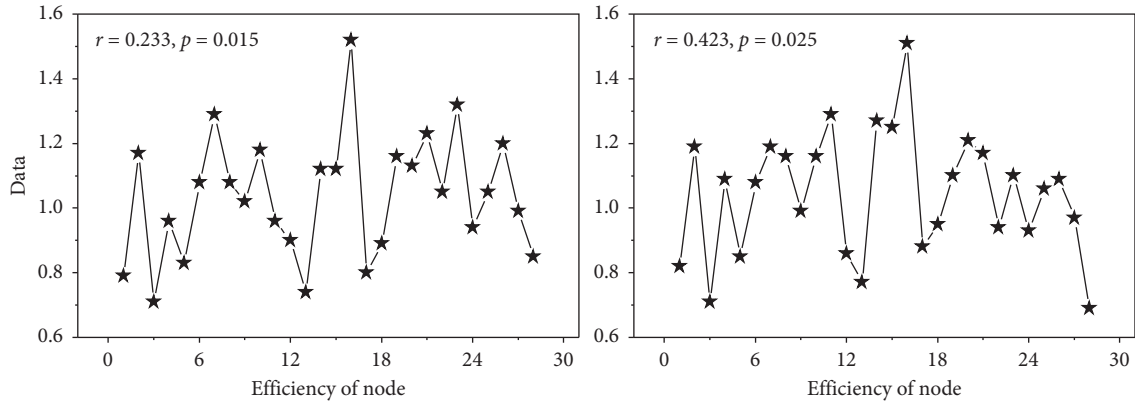


FIGURE 10: RNN residual distribution.

algorithm itself is easy to fall into local optimum, which is mainly due to the fast global search ability of PSO algorithm in parameter search and the fastest speed to select the best parameters.

The actual prediction of the three prediction algorithms is tested in detail here. For the RNN network, the test points on the training set are selected, and the RNN residual distribution is drawn as in Figure 10. After the analysis, it can be found that the first 200 training sample data can show the horizontal band pattern of the residuals well, but the corresponding error's allowed range is floating between -1 and 1 . This is due to the “forgetting” problem of the RNN prediction algorithm over a long period of time, and the overall error is found to be relatively large under RNN training. In terms of NSSP, a network security posture prediction algorithm based on PSO-LSTM neural network is proposed to address the problems of poor applicability and low accuracy of traditional network security posture prediction algorithms. Firstly, the recurrent neural network (RNN) is used to predict the security situation, and the LSTM neural network with “gating” structure is adopted to address the problems of insufficient memory of RNN, easy “forgetting” and gradient disappearance during the training process. The gradient disappearance was solved by reasonably controlling the ratio of input information to current memory information, and the training prediction results were significantly improved. However, it is difficult to select the parameters of the LSTM network during the training process, and it is easy to fall into the local optimum, so the particle swarm optimization algorithm is introduced to achieve the global optimum quickly, reduce the training time, and improve the efficiency. The accuracy and effectiveness of the PSO-LSTM prediction algorithm proposed in this paper are demonstrated through experimental analysis and comparison.

After the training of the PSO-LSTM neural network prediction algorithm, it was tested on the test set individually and the relative error of each sample was obtained. It is found that the accuracy of the proposed prediction algorithm in the test set is slightly lower than that in the training set, but the overall prediction accuracy is still acceptable, and the average relative error is stable at about 5.07% (only 30 samples are listed here). At the same time, the security

posture prediction curves on the test set are plotted, and the real values of the posture are compared with the predicted values of the PSO-LSTM algorithm, and the results show that the prediction algorithm proposed in this paper can predict the network security posture in the future period better. In terms of NSSE, a hierarchical situational assessment model based on alert verification and fuzzy inference is proposed to address the problem that the current network security situational assessment model only uses a large amount of IDS alert information without effectively combining the target system configuration information, asset value information, and vulnerability information. And the evaluation system indicators suitable for this model are determined; secondly, the alarm verification process based on multilayer fuzzy mathematical evaluation is implemented to reduce the influence of false alarms, make each alarm message more targeted, and obtain alarm success rate information; then, the fuzzy inference process is applied to achieve a nonlinear mapping of the three alarm elements to synthesize the alarm posture values; then, the alarm posture values are obtained through different levels at the service level; finally, the security posture values are quantitatively calculated at service level, host level, and network system level through different levels, and the threat status at each level is displayed in a hierarchical manner. The assessment model can obtain more comprehensive assessment information and more accurate assessment results than traditional methods.

4. Conclusion

In this paper, we focus on network security situation prediction and propose a PSO-LSTM neural network-based network security situation prediction algorithm. Particle swarm optimization (PSO) is an evolutionary computation technology (evolutionary computation). The basic idea of particle swarm optimization algorithm is to find the optimal solution through collaboration and information sharing between individuals in the group. In order to solve the problem of insufficient memory and the disappearance of the gradient of the recurrent neural network (RNN); the LSTM neural network with gating structure is used to reasonably control the ratio of input information to current

memory information, and the gradient disappears. LSTM (long short-term memory) is specially designed to solve long-term problems. All RNNs have a chain form of repeating neural network modules. In a standard RNN, this repeated structural module has only a very simple structure, such as a tanh layer. In order to solve the problem of difficult parameter selection of LSTM network and easy to fall into local optimum, particle swarm optimization algorithm is added in the network training to achieve global optimum quickly, reduce the training time, and improve the efficiency. Finally, the three prediction algorithms are compared by experimental analysis, which proves the accuracy and effectiveness of the PSO-LSTM prediction algorithm proposed in this paper, and its application in the field of situational prediction has good results.

Data Availability

No data were used to support this study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] H. Shakhatareh, A. H. Sawalmeh, A. Al-Fuqaha et al., "Unmanned aerial vehicles (UAVs): a survey on civil applications and key research challenges," *IEEE Access*, vol. 7, pp. 48572–48634, 2019.
- [2] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and L. Hanzo, "Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1656–1672, 2017.
- [3] Y. Meidan, M. Bohadana, Y. Mathov et al., "N-BaIoT-Network-Based detection of IoT botnet attacks using deep autoencoders," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 12–22, 2018.
- [4] Y. Wu, A. Khisti, C. Xiao, G. Caire, K.-K. Wong, and X. Gao, "A survey of physical layer security techniques for 5G wireless networks and challenges ahead," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 679–695, 2018.
- [5] N. C. Luong, D. T. Hoang, S. Gong et al., "Applications of deep reinforcement learning in communications and networking: a survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3133–3174, 2019.
- [6] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When intrusion detection meets blockchain technology: a review," *IEEE Access*, vol. 6, pp. 10179–10188, 2018.
- [7] D. Kwon, "A survey of deep learning-based network anomaly detection," *Cluster Computing*, vol. 22, no. 1, pp. 949–961, 2019.
- [8] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2595–2621, 2018.
- [9] A. Lei, H. Cruickshank, Y. Cao, P. Asuquo, C. P. A. Ogah, and Z. Sun, "Blockchain-based dynamic key management for heterogeneous intelligent transportation systems," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1832–1843, 2017.
- [10] L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis, "Secrecy cooperative networks with outdated relay selection over correlated fading channels," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7599–7603, 2017.
- [11] D. B. Rawat and S. R. Reddy, "Software defined networking architecture, security and energy efficiency: a survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 325–346, 2017.
- [12] Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 760–776, 2019.
- [13] B. He, A. Liu, N. Yang, and V. K. N. Lau, "On the design of secure non-orthogonal multiple access systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2196–2206, 2017.
- [14] R. I. Ansari, C. Chrysostomou, S. A. Hassan et al., "5G D2D networks: techniques, challenges, and future prospects," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3970–3984, 2018.
- [15] W. L. Al-Yaseen, Z. A. Othman, and M. Z. A. Nazri, "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system," *Expert Systems With Applications*, vol. 67, pp. 296–303, 2017.
- [16] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on SDN based network intrusion detection system using machine learning approaches," *Peer-to-Peer Networking and Applications*, vol. 12, no. 2, pp. 493–501, 2019.
- [17] J. Chen, L. Yang, and M.-S. Alouini, "Physical layer security for cooperative NOMA systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4645–4649, 2018.
- [18] M. Chen, Y. Tian, G. Fortino, J. Zhang, and I. Humar, "Cognitive Internet of vehicles," *Computer Communications*, vol. 120, pp. 58–70, 2018.
- [19] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2702–2733, 2019.
- [20] K. Kalkan and S. Zeadally, "Securing Internet of things with software defined networking," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 186–192, 2018.
- [21] N. Z. Bawany, J. A. Shamsi, and K. Salah, "DDoS attack detection and mitigation using SDN: methods, practices, and solutions," *Arabian Journal for Science and Engineering*, vol. 42, no. 2, pp. 425–441, 2017.
- [22] Z. Zhang, Y. Yu, H. Zhang et al., "An overview of security support in named data networking," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 62–68, 2018.
- [23] C. C. Byers, "Architectural imperatives for fog computing: use cases, requirements, and architectural techniques for fog-enabled IoT networks," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 14–20, 2017.
- [24] K. Cabaj, "Software-defined networking-based crypto ransomware detection using HTTP traffic characteristics," *Computers & Electrical Engineering*, vol. 66, pp. 353–368, 2017.
- [25] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "Towards detecting compromised accounts on social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 4, pp. 447–460, 2017.

Research Article

ITDPM: An Internet Topology Dynamic Propagation Model Based on Generative Adversarial Learning

Hangyu Hu , Xuemeng Zhai , Gaolei Fei, and Guangmin Hu

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Correspondence should be addressed to Xuemeng Zhai; zxm@uestc.edu.cn

Received 26 April 2021; Accepted 19 May 2021; Published 29 May 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Hangyu Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Network information propagation analysis is gaining a more important role in network vulnerability analysis domain for preventing potential risks and threats. Identifying the influential source nodes is one of the most important problems to analyze information propagation. Traditional methods mainly focus on extracting nodes that have high degrees or local clustering coefficients. However, these nodes are not necessarily the high influential nodes in many real-world complex networks. Therefore, we propose a novel method for detecting high influential nodes based on Internet Topology Dynamic Propagation Model (ITDPM). The model consists of two processing stages: the generator and the discriminator like the generative adversarial networks (GANs). The generator stage generates the optimal source-driven nodes based on the improved network control theory and node importance characteristics, while the discriminator stage trains the information propagation process and feeds back the outputs to the generator for performing iterative optimization. Based on the generative adversarial learning, the optimal source-driven nodes are then updated in each step via network information dynamic propagation. We apply our method to random-generated complex network data and real network data; the experimental results show that our model has notable performance on identifying the most influential nodes during network operation.

1. Introduction

Nowadays, from various telecommunication systems to power grid systems, it can be seen that everyone's lives are affected and dominated by today's real-world complex networks [1–5]. However, although various emerging network technologies have brought about more and more convenience to humans in many fields, the network is also vulnerable to potential risks and threats. The research of the information propagation such as spreading of rumors, influence diffusion, packet forwarding, and epidemic proportions has a long tradition in network science including social network analysis, Internet topology analysis, and complex biological network analysis. Identifying the influential source nodes that make the information propagates as quickly as possible is one of the most important problems in

network science, and many researchers did hard works on finding such influential nodes [6–11].

With the classic infectious disease models including SI, SIS, and SIR model proposed, researchers began to pay attention to the impact of the network structure on the propagation behavior. However, most analyses of the propagation are directed to macroscopic statistical characteristics of network structures, ignoring the actual connection of links. Research on influential nodes' identification has to be based on the clear network structure, and the optimal influence problem is shown to be NP-hard [12]. Artificial intelligence emerging in recent years aims to solve the difficult problems using machine learning such as the generative adversarial network (GAN) [13]. The ideal of the GAN is to optimize the results of the generator through constant confrontation between the generator and the

discriminator. The process of generator optimization could be regarded as a kind of generative adversarial learning [14, 15]. Therefore, with the mature complex network theory and new artificial intelligence methods, the problems of the influential nodes' identification could be solved efficiently.

In this paper, we propose an Internet Topology Dynamic Propagation Model (ITDPM) to find the influential nodes in the Internet topology. The model consists of two parts: the generator and the discriminator like the GANs. The generator aims to identify the influential nodes that make the information propagate as quickly as possible in the Internet topology at the current stage. Two famous theory methods in network science are used and modified to be adaptive to the characteristics of the Internet topology: network controllability and node importance parameters. The discriminator is designed to simulate the information propagation and evaluate the performance based on the influential nodes that the generator generates. The performance of the propagation will be fed back to the generator to help the model update the set of influential nodes. The propagation rate and the coverage are used as the general parameters to evaluate the performance of the generator and determine the stopping conditions of the model. Experiments on routing attack propagation are conducted based on both simulated and real-world network topology to show the performance of ITDPM on the influential nodes' identification. The experimental results show notable performance in identifying the optimal influential nodes to make the information propagate as quickly as possible through our model.

The main contributions of the present research are summarized as follows:

- (1) According to the minimum input theorem, we can discover the minimum set of driving nodes in the Internet topology to ensure that the speed and breadth of information propagation from these nodes can reach the fastest speed.
- (2) This paper utilizes ITDPM as an effective tool to identify the optimal influential nodes in Internet topology. The model consists of two processing stages: the generator and the discriminator like the GANs.
- (3) The experimental results show that, compared to traditional methods, our approach effectively identifies the optimal influential nodes in Internet topology and is also useful for improving network security management by enhancing critical nodes' security policy.

The rest of this paper is organized as follows: Section 2 briefly surveys the related work of network vulnerability analysis and GAN. Section 3 introduces the preliminaries and problem definition of critical node detection in Internet topology. The framework of using the ITDPM for the optimal influential node identification is discussed in Section 4. Section 5 presents the experiments on information propagation and simulated routing attack. Finally, Section 6 gives the conclusions and presents possible future work.

2. Related Work

The study on epidemics offers powerful models for analyzing the information propagation, such as the SI, SIS, and SIR model [16, 17]. However, those models do not care about the specific structure of the networks. Therefore, the problem of influential nodes' identification based on the actual structure of the networks has become a significant issue in recent years. In [6], the problem to find the minimal set of influential nodes was mapped onto optimal percolation in random networks, which were arisen by minimizing the energy of a many-body system, where the form of the interactions is fixed by the nonbacktracking matrix of the network. In [18], the influential nodes were identified through the family of H-indices. The convergence to coreness could be guaranteed even under an asynchronous updating process, allowing a decentralized local method to calculate a node's coreness in large-scale evolving networks. In [19], the optimal percolation theory was used to predict and subsequently target nodes that are essential for global integration of a memory network in rodents. In [20], a node information dimension is proposed by synthesizing the local dimensions at different topological distance scales. In [21], the spreading capacity of the focal node was accurately characterized by assigning different weights for each class of neighbors and summing up the neighbors' contributions.

However, the GAN in the field of artificial intelligence is widely used to solve difficult problems in various disciplines [22]. In [23], the GAN was used to predict socially plausible futures and encourage diverse predictions with a novel variety loss. In [24], a novel method to fuse two types of information using a GAN was proposed, termed as FusionGAN. In [25], a new GAN-based model was presented to calculate for each large transfer probability that it is fraudulent, such that the bank can take appropriate measures to prevent potential fraudsters from taking the money if the probability exceeds a threshold. To the best of our knowledge, few works of the GAN are applied on the Internet topology or the complex structure.

3. Problem Definition

In the real-world Internet topology, network packets are delivered based on the specific routing protocol, such as BGP, TCP, and ICMP. The forwarding path of the packets can be used as propagation path of the information because the packets themselves can be the carriers of network information propagation or some malicious attacks. Therefore, we can model the forward path of the Internet topology to form a complex network with nodes and edges. Figure 1 shows the routing example of the real word by the CAIDA research institute [26]. The definition is shown as follows.

Definition 1. Internet Routing Propagation Network. Given a directed network $G = \langle V, E \rangle$, where V refers to the set of nodes and E refers to the set of directed edges among nodes, the routers and hosts are abstracted as the nodes in G and the forwarding path for router i to router j is abstracted as the

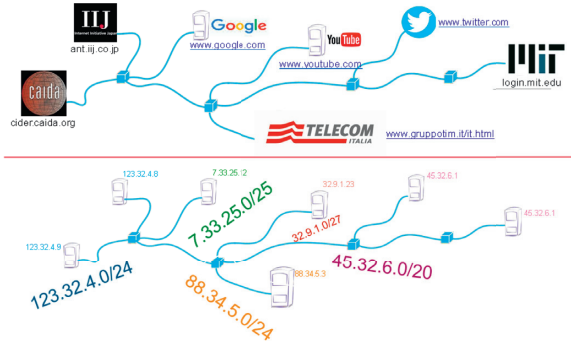


FIGURE 1: The routing example of the real word by CAIDA research institute.

edge e_{ij} meaning that node i points to the node j . The directed network G denotes the Internet routing propagation network.

Figure 2 shows a graph model example of the Internet routing propagation network built based on the data from CAIDA [27]. The hosts are hidden for the visualization. Though the real-world forwarding paths based on the Internet topology are abstracted as the theoretical network, they cannot be processed directly by the machine or participate in computing operations. In network science, the adjacency matrix addresses the problem as a two-dimension matrix to represent the network. Therefore, we use the adjacency matrix to describe the directed Internet routing propagation network, denoted as follows.

Definition 2. Transfer Adjacency Matrix. Given an adjacency matrix $A_{N \times N}$ of directed network G , where N refers to the number of nodes in G , the elements $a_{ij} = 1$ if there is an edge pointing from node i to node j and otherwise $a_{ij} = 0$. The adjacency matrix $A_{N \times N}$ is donated as the transfer adjacency matrix only if it is the adjacency matrix of the Internet routing propagation network. The element a_{ij} refers to the packets forwarding path between the routers or the hosts.

The Internet routing propagation network could be abstracted as a complex network and be represented through the transfer adjacency matrix. Our goal is to find the source nodes that make the information propagates as quickly as possible. Therefore, a propagation model has to be built to simulate and analyze the propagation process. Since the router protocol is updated regularly during a specific interval, ideally, we assume that information is propagated through each update. Therefore, our main problems could be summarized as follows.

Problem 1. Propagation Model Construction. Given the Internet routing propagation network G and its transfer adjacency matrix A , the first problem is to build a propagation model to simulate the propagation process of the information. The information is assumed to propagate through each specific interval of the routing update.

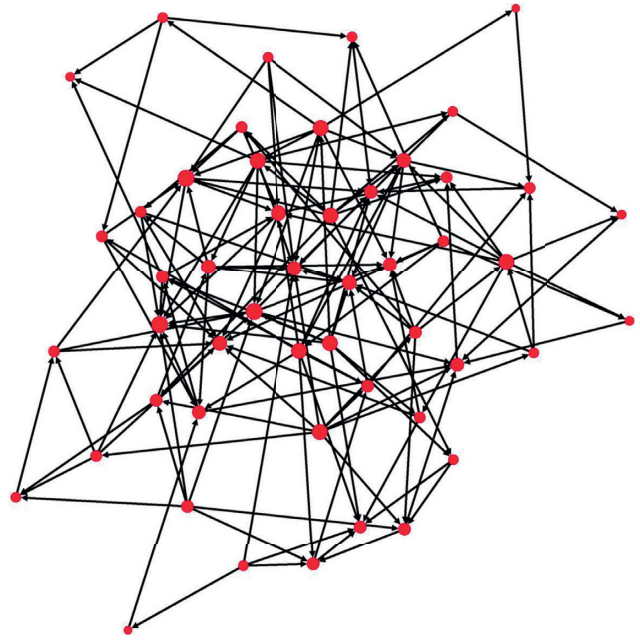


FIGURE 2: The graph model of the Internet routing propagation network.

Problem 2. Influential Nodes' Identification. Given the propagation model, the problem is to find a set of optimal source nodes that make the information propagate as quickly as possible. Such a set of source nodes is denoted as the set of influential nodes R .

With the two problems, we build the ITDPM as the propagation model M to identify the influential nodes R . The detailed model is discussed in the next section.

4. Internet Topology Dynamic Propagation Model

In this section, a detailed description of our ITDPM is given. The main idea of the ITDPM is based on generative adversarial learning. It consists of two parts, including the generator and the discriminator like the GANs. However, there is no network in the model. Therefore, the generation and discrimination process could be regarded as a kind of generative adversarial learning. The components of ITDPM are shown in Figure 3.

The two main components of ITDPM are the generator and discriminator. The task of the generator is to analyze the Internet routing propagation network and identify the influential nodes that make the information propagate as quickly as possible. The methods of influential nodes' identification combine the network controllability and node importance parameter in the complex network. Both theories have been modified to be adaptive to the characteristics of the Internet topology. The discriminator aims to simulate the information propagation process and compute the spreading parameters to evaluate the performance of the generator. The remaining topology will be fed back to the

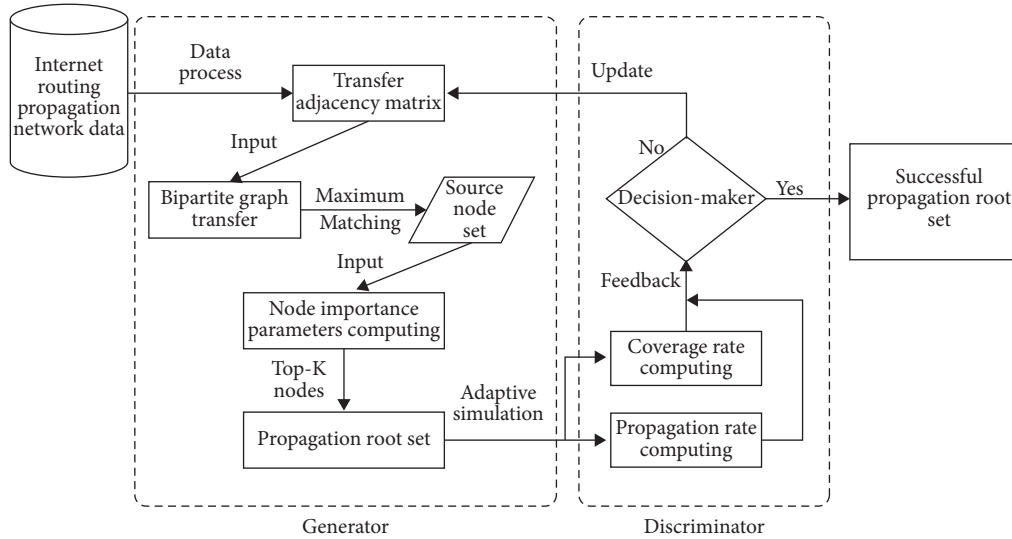


FIGURE 3: The system components of the ITDPM.

generator to find other better source nodes. There are two situations that cause the feedback: (1) the performance dissatisfies the requirements including slow propagation speed and low coverage; (2) the propagation stops due to the structure of the directed network. These two situations will lead to the stagnation of information dissemination in the Internet.

The ITDPM gets the optimal influential nodes R after several times of generating and discriminating. The outputs of the generator are the influential nodes in each propagation stage. The outputs of the discriminator are the parameters that could evaluate the performance of the information propagation based on the influential nodes from the generator. Two general parameters are used in the ITDPM: the propagation rate curve and the coverage. Other discriminant parameters could be added to the discriminator to evaluate the performance of the generator and determine the stopping conditions of the model, such as the propagation delay and packet loss rate in the routing attack propagation network. The ITDPM will stop and output the final optimal influential nodes.

4.1. The Generator Stage. The generator aims to identify the influential nodes in the Internet routing propagation network. Given a directed Internet routing propagation network G , the primary task is to identify the source routers for the packet forwarding. The second task is to filter the optimal nodes from the source routers that could make the information propagate as quickly as possible. Therefore, two well-known theoretical methods in network science are used and modified to be adaptive to the characteristics of the Internet topology: network controllability and node importance parameters. We first give the details of the two theories in the generator.

4.1.1. Network Controllability. The network controllability is also known as the structural controllability [28]. In the most real-world network system, the topology structure is the only

known condition, and the weights of the connection are usually unknown. The structural controllability is to study how to identify the minimum number of driver nodes to control the whole network under the unknown strength of the interaction among nodes. Therefore, the minimum input theorem is proposed to get the driver nodes. It is proved that, in order to fully control a directed network G , the minimum number of input nodes (or equivalently the minimum number of driver nodes) is related to the size of a maximum matching in G . Therefore, we first introduce the concept of the network matching.

4.1.2. Network Matching. M is an independent edge set without common nodes. A node is matched if it is incident to an edge in the matching. For a directed network, an edge subset M is matching if no two edges in M share a common starting node or a common ending node. A node is matched if it is an ending node of an edge in the matching.

In both networks, the matching of maximum size is called maximum matching. In general, there could be many different maximum matchings for a given network. A maximum matching is called perfect if all nodes are matched. Therefore, to analyze the structural controllability of the network, the directed network G needs to be transferred to the bipartite graph set B . The bipartite graph contains two sets of nodes: the set of starting nodes S and the set of end nodes D . The edges only exist between the nodes in different two sets. The unmatched nodes and matched starting nodes in the maximum matching of the bipartite graph are the driver nodes. The process of bipartite graph transfer and matching is shown in Figure 4.

The driver node set contains two categories: the source nodes of the propagation links and the nodes on the branch links. The source nodes of the propagation links are our target nodes. Therefore, the minimum input theorem is changed to only get the root nodes in the maximum matching of the bipartite graph shown in Figure 4. It is easy to find the root nodes because each matching edge consists

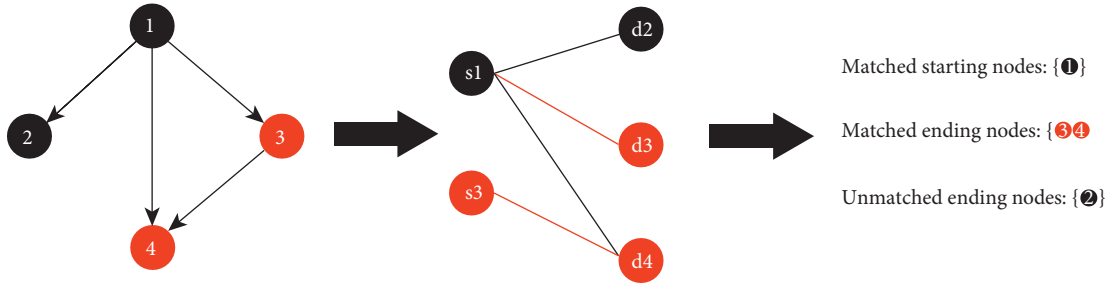


FIGURE 4: The process of the bipartite graph transfer and matching.

of the starting node and the end node. The root node of the link could be traced back by finding the matching between the starting node and the ending node of the two edges.

The root nodes found through the modified minimum input theorem are the source nodes of the directed links in a directed network. In the Internet routing propagation network, those nodes are the starting nodes of the packet forwarding. In other words, they are the source nodes of information propagation, denoted as set S .

4.1.3. Node Importance Parameter. With the development of network science, several parameters that describe the importance of nodes are proposed in the past decades. In this paper, two node importance parameters that best match the characteristics of the Internet topology are used to measure the influence of the source nodes S . The influence here refers to the impact of the source nodes on the information propagation rate and coverage of the entire Internet topology. These node importance parameters are the betweenness centrality and k -shell. The basic conception of the two parameters is as follows.

Betweenness centrality is a parameter of the influence measurement of a node over the links of information between every pair of nodes under the assumption that information primarily propagates over the shortest paths between them, denoted as follows:

$$BC_i = \sum_{s \neq i \neq t} \frac{n_{st}^i}{g_{st}}, \quad (1)$$

where g_{st} refers to the number of shortest paths from node s to node t and n_{st}^i refers to the number of all shortest paths from node s to node t through node i .

k -shell decomposition is an extension of the importance ranking of nodes based on the node degree. The process of k -shell decomposition is to iteratively remove the nodes of which degree equals k until there is no such node in the network from $k = 1$. The k -shell nodes consist of the removed nodes at k times.

Figure 5 shows the betweenness centrality and k -shell. The betweenness centrality is consistent with the requirements of most routing protocols such as RIP and OSPF for the shortest path propagation. The k -shell matches the hierarchical structure of the Internet topology. The nodes with large k -shell tend to be the core routers in the Internet topology. Therefore, the information carried by the packet

forwarding can be propagated faster from nodes with both high betweenness centrality and k -shell.

In order to measure the influence of the nodes in the set of source nodes S , the normalization equation of the two node importance parameters is used to describe such influence I , denoted as follows:

$$I_s = \alpha * BC_i + \beta * k_{i\text{-shell}}, \quad (2)$$

where α and β are the coefficients that balance the weights of the two parameters.

4.1.4. The Process of the Generator Stage. When computing the influence I_s of each source node s in S , the i -step influential nodes R_i could be defined as the top- k influence source nodes in the set of source nodes S . At each step of the propagation, the current optimal influential nodes are generated through the improved methods in network science and transferred to the discriminator to get the simulated performance of the propagation. The influential nodes are updated after the discriminator sends a new network topology.

4.2. The Discriminator Stage. The main task of the discriminator is to simulate the information propagation and evaluate the performance based on the influential nodes that the generator generates. The performance of the propagation will be fed back to the generator to help the model update the set of influential nodes. Our ITDPM assumes that the information is propagated through each update of the router protocol during a specific interval. Therefore, each interval could be regarded as a time unit of propagation. The information is propagated from one node to another over a time unit. With the ideal hypothesis, it is easy to simulate and evaluate the performance of the propagation process. We first introduce the two general parameters for the propagation evaluation.

4.2.1. Evaluation. The two parameters are the propagation rate and the coverage rate. Both parameters only require the basic abstracted nodes and propagation links. They do not require any other specific requirements such as node attributes and edge weights. Therefore, the two parameters are the general measurements for the evaluation.

Propagation rate refers to the number of nodes that are propagated per time unit, denoted as

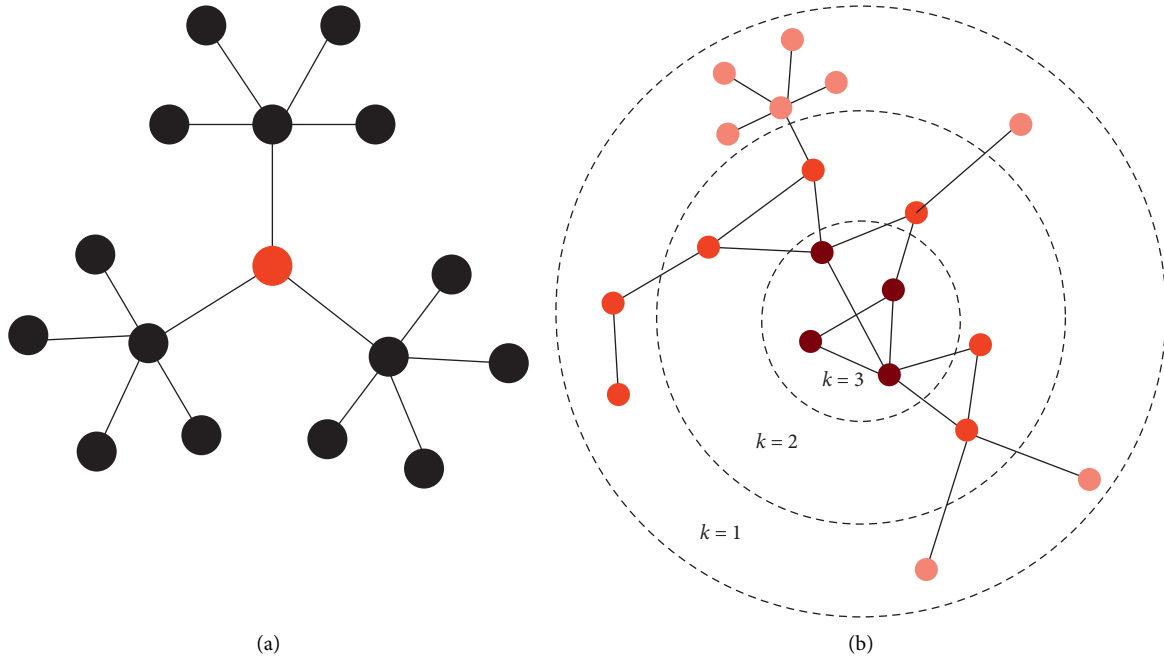


FIGURE 5: The schematic diagram of betweenness centrality and (k) -shell: (a) the red node is with the highest betweenness centrality; (b) the nodes with darker color refer to the larger (k) -shell.

$$v_r = \frac{dN_p}{dt}, \quad (3)$$

where N_p is the number of nodes that are propagated. The average propagation rate $v_a = N_p/t$.

Coverage rate refers to the ratio of the number of propagated nodes to the total number of nodes in the network, denoted as

$$c_r = \frac{N_p}{N}. \quad (4)$$

The propagation rate describes the speed of the propagation, and the coverage rate stands for the breadth of the propagation. The two parameters comprehensively evaluate the propagation process in a general sense. To some specific situation, some other parameters could be used to evaluate the performance such as the propagation delay and packet loss rate in the routing attack propagation.

4.2.2. The Process of the Discriminator Stage. The discriminator will determine whether it is a successful propagation according to the value of the propagation rate v_r and coverage rate c_r . If both v_r and c_r are larger than the specified thresholds, the decision parameter D will be 1 to show that it is a successful propagation. Otherwise, the decision parameter D will be 0, and the generator will update the propagation roots based on the remaining network structure which consists of those nodes not propagated.

5. Experimental Results

In order to show the performance of our ITDPM on the influential nodes' identification, experiments on the routing attack propagation are conducted based on both simulated and real-world network topologies. It is assumed that the attack could cause the routers unworkable and the attack will be carried on the updated packets of the routing protocol. The simulated network topology is generated by NS-3 simulator with 3,000 nodes and 4,513 directed edges. The forwarding paths are computed through the router table of each router in the simulated network topology. The real-world network topology is generated based on the data from SNAP. It is the AS Internet topology from CAIDA and contains 26,475 nodes and 27,562 edges. Table 1 shows the detailed topological parameters of simulated data and real-world data. Both networks are directed. The simulated network topology is like a local network topology of the small scale and the AS Internet topology is the global networks of a large scale.

Two methods are compared with our ITDPM to identify the influential nodes that make the attack propagate as quickly as possible: the random identification (RI) and the maximum out-degree identification (MDI). The RI selects n nodes randomly as the root nodes to propagate the attack information. The MDI is to select the top- n nodes with the largest out-degree. The number of influential nodes is limited to 0.5% of the total number of nodes in the network topology. Therefore, in the simulated network topology, n is

TABLE 1: Topological parameters of simulated data and real-world data.

Name	Number of nodes	Number of edges	Avg. deg.	Diameter	K-shell layers
NS-3 data	3,000	4,513	2.75	15	3
CAIDA data	26,475	27,562	2.08	18	5

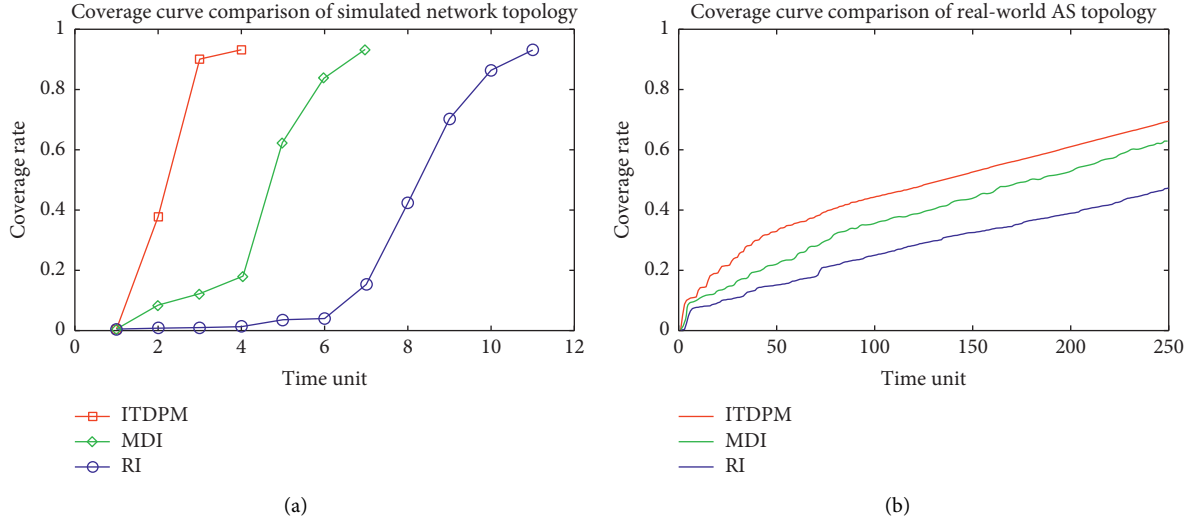


FIGURE 6: The comparison results of the coverage curve in both simulated network topology and real-world AS topology.

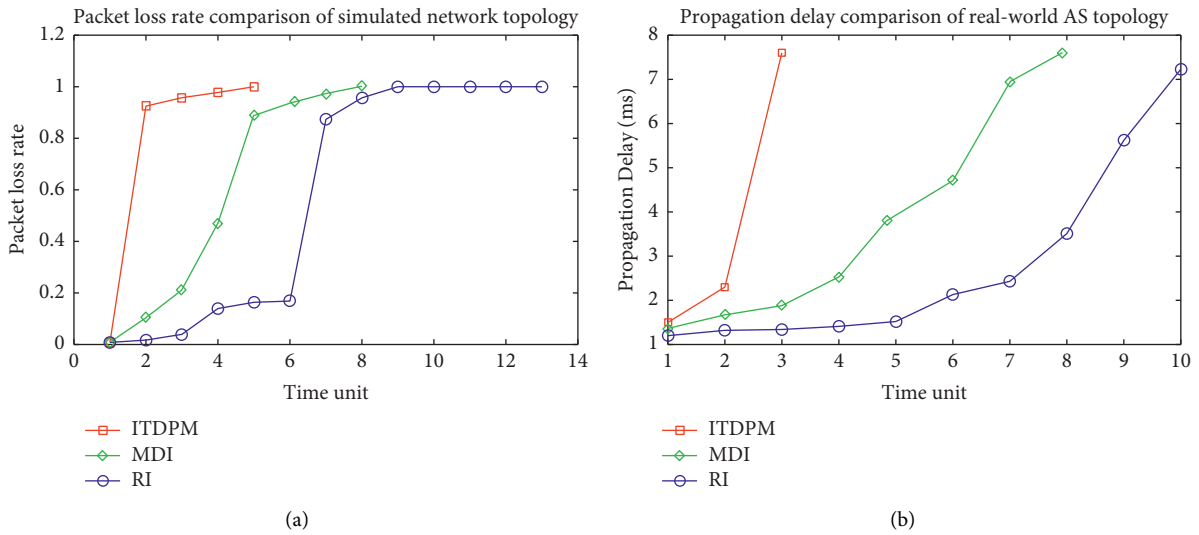


FIGURE 7: The comparison results of packet loss rate and propagation delay in simulated network topology and real-world topology.

no larger than 15, and in the real-world AS topology, n is smaller than 130. To make the two methods get the most optimal influential nodes, the value of n is selected as the upper bound of the range.

The coverage curve is used to evaluate the performance of the propagation. The coverage curve describes the coverage of propagated nodes at time unit. The derivative of the curve is the propagation rate. The curve can be computed in both simulated network topology and real-world AS topology. Particularly, in the simulated environment, two parameters could be computed to describe the influence of

the attack: packet loss rate and propagation delay. The two parameters could describe the degree of damage to the network topology. Therefore, two parameters are computed to evaluate the simulated network topology as the additional measures.

The comparison results of the coverage curves in both simulated network topology and real-world AS topology are shown in Figure 6. Both results show the better performance of our method. In the simulated network topology with 1,000 nodes, the attack could be propagated nearly all nodes in 4 time units with our method and in 7 and 11 time units,

respectively, with the MDI and RI. In the small-scale network, it only needs one iteration for our ITDPM to find the optimal influential nodes, which demonstrates the effectiveness of the algorithms in the generator. In the real-world AS topology with 26,475 nodes, after 100 time units, the coverage rate of all methods is nearly linear. Before 100 time units, the result of ITDPM keeps the fast propagation rate. Each jump of the red curve represents the antagonism between the generator and the discriminator. The propagation rate will jump after the generator updates the set of the influential nodes.

The comparison results of the packet loss rate and the propagation delay in simulated network topology are shown in Figure 7. The two parameters evaluate the performance of the attack. As shown in experiment results, our ITDPM shows the best performance among all three methods. The packet loss rate and propagation delay show a strong correlation with the coverage rate shown in Figure 6. The packet loss of ITDPM reaches a quite high point at the second time unit, which means that the attack based on our ITDPM breaks the topology fast and completely. After 4 time units, all packets are lost under nearly 100% coverage rate of the propagated nodes. With the increase of the packet loss rate and coverage rate, the propagation delay based on our ITDPM grows exponentially to very high at the third time unit. After that, it becomes positive infinity because most of the nodes are unreachable.

6. Conclusion and Prospction

The results of both general parameters and specific parameters show a notable preference on the information propagation analysis and influential nodes' identification of our ITDPM. The rapid coverage rate demonstrates that our method can find the source nodes that make the information propagate as quickly as possible. The process of the antagonism between the generator and the discriminator could update the influential nodes based on the current network topology structure to help the information propagate faster. Our ITDPM works effectively on the attack propagation of the Internet routing topology based on the results of the packet loss rate and the propagation delay. The specific parameters indicate that the Internet topology is effectively broken based on our ITDPM.

In general, our ITDPM is a theoretical model of propagation analysis and influential nodes' identification based on generative adversarial learning. It can be used for any propagation problem based on the practical network structure, such as the spread of influence and rumors in social networks. The basic task for our model is to find the influential nodes that make the information propagate as quickly as possible. Our future work is dedicated to improving both generator and discriminator to get a more effective model in information propagation analysis.

Data Availability

The data sets used to obtain the results in this manuscript are available at <http://snap.stanford.edu/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 62071095) and the Sichuan Science and Technology Program under Grant 2019YFG0456.

References

- [1] J. Su, Z. Sheng, A. X. Liu, Y. Han, and Y. Chen, "Capture-aware identification of mobile RFID tags with unreliable channels," *IEEE Transactions on Mobile Computing*, vol. 1, 2020 (in press).
- [2] J. Su, A. X. Liu, Z. Sheng, and Y. Chen, "A partitioning approach to RFID identification," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2160–2173, 2020.
- [3] J. Su, R. Xu, S. Yu, B. Wang, and J. Wang, "Idle slots skipped mechanism based tag identification algorithm with enhanced collision detection," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 5, pp. 2294–2309, 2020.
- [4] J. Su, R. Xu, S. Yu, B. Wang, and J. Wang, "Redundant rule detection for software-defined networking," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 2735–2751, 2020.
- [5] J. Su, Z. Sheng, A. X. Liu, Z. Fu, and Y. Chen, "A time and energy saving-based frame adjustment strategy (TES-FAS) tag identification algorithm for UHF RFID systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 2974–2986, 2020.
- [6] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," *Nature*, vol. 524, no. 7563, pp. 65–68, 2015.
- [7] W. Chen, M. Jiang, C. Jiang et al., "Critical node detection problem for complex network in undirected weighted networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 538, 2020.
- [8] M. Lalou, M. A. Tahraoui, and H. Kheddouci, "The critical node detection problem in networks: a survey," *Computer Science Review*, vol. 28, pp. 92–117, 2018.
- [9] D. Santos, A. de Sousa, and P. Monteiro, "Compact models for critical node detection in telecommunication networks," *Electronic Notes in Discrete Mathematics*, vol. 64, pp. 325–334, 2018.
- [10] G. Alozie, A. Arulselman, K. Akartunali et al., "Efficient methods for the distance-based critical node detection problem in complex networks," *SSRN Electronic Journal*, vol. 131, 2021.
- [11] M. Ficco, M. Choraś, and R. Kozik, "Simulation platform for cyber-security and vulnerability analysis of critical infrastructures," *Journal of Computational Science*, vol. 22, pp. 179–186, 2017.
- [12] D. Kempe, J. Kleinberg, and É Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146, Washington, DC, USA, August 2003.
- [13] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza et al., *Generative adversarial networks*, 2014.
- [14] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial

- networks: an overview,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 53–65, 2018.
- [15] J. Ho and S. Ermon, “Generative adversarial imitation learning,” in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, pp. 4572–4580, Barcelona, Spain, December 2016.
- [16] R. Pastor-Satorras and A. Vespignani, “Epidemic spreading in scale-free networks,” *Physical Review Letters*, vol. 86, no. 14, pp. 3200–3203, 2001.
- [17] M. Boguná and R. Pastor-Satorras, “Epidemic spreading in correlated complex networks,” *Physical Review E*, vol. 66, no. 4, Article ID 047104, 2002.
- [18] L. Lü, T. Zhou, Q. M. Zhang et al., “The H-index of a network node and its relation to degree and coreness,” *Nature Communications*, vol. 7, no. 1, pp. 1–7, 2016.
- [19] G. Del Ferraro, A. Moreno, B. Min et al., “Finding influential nodes for integration in brain networks using optimal percolation theory,” *Nature Communications*, vol. 9, no. 1, pp. 1–12, 2018.
- [20] T. Bian and Y. Deng, “Identifying influential nodes in complex networks: a node information dimension approach,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 4, Article ID 043109, 2018.
- [21] C. Li, L. Wang, S. Sun, and C. Xia, “Identification of influential spreaders based on classified neighbors in real-world complex networks,” *Applied Mathematics and Computation*, vol. 320, pp. 512–523, 2018.
- [22] X. Wu, K. Xu, and P. Hall, “A survey of image synthesis and editing with generative adversarial networks,” *Tsinghua Science and Technology*, vol. 22, no. 6, pp. 660–674, 2017.
- [23] A. Gupta, J. Johnson, L. Fei-Fei et al., “Social gan: socially acceptable trajectories with generative adversarial networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2255–2264, Salt Lake, UT, USA, June 2018.
- [24] J. Hou, D. Zhang, W. Wu, J. Ma, and H. Zhou, “A generative adversarial network for infrared and visible image fusion based on semantic segmentation,” *Entropy*, vol. 23, no. 3, 376 pages, 2021.
- [25] Y.-J. Zheng, X.-H. Zhou, W.-G. Sheng, Y. Xue, and S.-Y. Chen, “Generative adversarial network based telecom fraud detection at the receiving bank,” *Neural Networks*, vol. 102, pp. 78–86, 2018.
- [26] D. Clark and K. Claffy: Knowledge of Internet Structure: Measurement, Epistemology, and Technology <https://www.caida.org/projects/kismet/>.
- [27] K. Claffy: Center for applied internet data analysis, <http://www.caida.org/home/>.
- [28] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, “Controllability of complex networks,” *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

Research Article

A Method of Amino Acid Terahertz Spectrum Recognition Based on the Convolutional Neural Network and Bidirectional Gated Recurrent Network Model

Tao Li ¹, Yuanyuan Xu ², Jiliang Luo,² Jianan He,³ and Shiming Lin ⁴

¹School of Mechanical Engineering, Hebei University of Technology, Tianjin 300401, China

²College of Information Science and Engineering, Huaqiao University, Xiamen 361021, China

³Central Laboratory of Health Quarantine,

Shenzhen International Travel Health Care Center and Shenzhen Academy of Inspection and Quarantine,
Shenzhen Customs District, Shenzhen 518033, China

⁴School of Informatics, Xiamen University, Xiamen 361005, China

Correspondence should be addressed to Yuanyuan Xu; yyxu@hqu.edu.cn and Shiming Lin; xmulsm@xmu.edu.cn

Received 22 April 2021; Accepted 24 May 2021; Published 29 May 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Tao Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the accuracy of amino acid identification, a model based on the convolutional neural network (CNN) and bidirectional gated recurrent network (BiGRU) is proposed for terahertz spectrum identification of amino acids. First, we use the CNN to extract the feature information of the terahertz spectrum; then, we use the BiGRU to process the feature vector of the amino acid time-domain spectrum, describe the time series dynamic change information, and finally achieve amino acid identification through the fully connected network. Experiments are carried out on the terahertz spectra of various amino acids. The experimental results show that the CNN-BiGRU model proposed in this study can effectively realize the terahertz spectrum identification of amino acids and will provide a new and effective analysis method for the identification of amino acids by terahertz spectroscopy technology.

1. Introduction

Terahertz (THz) waves are electromagnetic waves spanning frequencies between 0.1 THz and 10 THz, occupying most of the electromagnetic spectrum between microwaves and infrared light waves [1]. The combination of the vibration and rotation frequency transition of biomolecules is in the terahertz band, and thus, the terahertz spectrum can reflect the subtle changes in molecular species and structure; this is called molecular fingerprinting [2]. The energy of terahertz photons is low, and the photon energy of the electromagnetic wave with a frequency of 1 THz can be set to about 4.1 meV, which can directly detect biological samples without destroying their structure. Compared with other detection methods, terahertz spectroscopy can realize the label-free, fast, and nondestructive detection of biomolecules

[3]. Therefore, it has great application potential in the field of biomedicine, and terahertz radiation has recently been used to study DNA, RNA [4, 5], amino acids [6], proteins [7], and other biomolecules [8].

Amino acid is the basic building blocks of proteins, and various proteins in living organisms are composed of 20 basic amino acids. Amino acids can be used not only as a marker for diseases but also as a therapeutic drug. Amino acids are representative biomolecules, and their rapid nondestructive detection and quantitative analysis are particularly important [9].

The terahertz fingerprint spectra of a large number of substances are superimposed on each other, which make the qualitative and quantitative analyses based on the spectra extremely difficult. Researchers mainly use machine learning or multivariate analysis (stoichiometry) to achieve

qualitative and quantitative identifications of biological samples [10]. Such methods include multiple linear regression (MLR), principal component analysis, and partial least squares [11, 12]. Ueno et al. [13] conducted a quantitative analysis on a mixture of different amino acids in 2006. In 2016, Lu et al. used partial least squares (PLS) and interval partial least squares (IPLS) regression to quantitatively analyze binary amino acids [9]. In order to improve the accuracy of quantitative analysis, some research groups have proposed the use of machine learning methods [10, 14]. Yuan et al. [14] performed a spectral classification of three fluoroquinolones based on the back propagation neural network (BPNN) and obtained an accuracy rate of 80.56%. Meanwhile, Peng et al. [15] used wavelet filtering combined with support vector machines to quantitatively analyze the main components of brain tissue; the root mean square error of this approach was 0.4%. Furthermore, Liu et al. [16] used the random forest (RF) algorithm to distinguish genetically modified rice seeds from nontransgenic rice seeds, and the classification accuracy of their model reached 96.67%.

Methods based on machine learning have a common problem, the need to manually extract features. This process is complicated and cumbersome, and some methods have limited nonlinear fitting capabilities, resulting in the lack of hidden information in the extracted features. In contrast, deep learning is very suitable for terahertz signal recognition, and they do not need manual extraction of features.

Among the deep learning algorithms, the bidirectional gated recurrent network (BiGRU) is a type of the bidirectional recurrent neural network model that can fully express the relationship between the current output of a sequence and previous information [17]. However, the characteristic dimension of the amino acid time series is too high, and the BiGRU is directly used to process the amino acid sequence parameters, which results in low efficiency. In addition, the convolutional neural network (CNN) has powerful feature extraction capabilities. To a certain extent, the more the layers, the more advanced the extracted features, the more the information contained, and the better the final classification result. In addition, the CNN requires fewer hyperparameters, has low computational complexity, and is widely used in image processing, semantic segmentation, and recognition.

Based on this, the present study uses the representative CNN and BiGRU to establish a CNN-BiGRU recognition model and makes full use of the advantages of the two networks to classify amino acids. First of all, we use the CNN to extract the abstract features of amino acid time series and then use BiGRU's dynamic timing information modeling ability to process the feature vectors of amino acid sequences.

Finally, experiments are carried out on the terahertz spectra of various amino acids. The experimental results show that the CNN-BiGRU model proposed in this study effectively realizes the terahertz spectrum identification of amino acids and avoids the tedious steps of artificial feature extraction, selection, and dimensionality reduction. Moreover, it demonstrates itself as a suitable terahertz spectrum technology for the identification of amino acids.

2. Methodology

2.1. CNN. Deep learning can ensure effective information extraction and feature expression and can complete tasks such as image recognition, time series prediction, and text prediction. Typical deep learning networks include the CNN, deep belief network, and recurrent neural network. Among them, the CNN can automatically learn filters and has developed into a mature feature extractor.

The CNN was proposed by Lecun et al. [18], and it is a feedforward multilayer neural network. It uses convolution operations to greatly reduce the dimensionality of the data and can achieve abstract expression of the original data [19, 20]. The basic structure of a CNN includes input, convolution, pooling, fully connected, and output layers, as shown in Figure 1.

The essence of the CNN is to construct multiple filters that can extract data features and the topological structure features hidden between the data using layer-by-layer convolution and pooling operations on the input data. Finally, these abstract features are merged through a fully connected layer, and the classification problem is solved through a Softmax or sigmoid activation function [21].

The convolution layer convolves the information in the receptive field by designing a convolution kernel of a suitable size and abstractly expresses the original data. When inputting data X , the feature map of the convolutional layer can be expressed as follows:

$$C = f(X \otimes W + b), \quad (1)$$

where \otimes is the convolution operation, W is the weight vector of the convolution kernel, b is the offset, and $f(\cdot)$ is the activation function, which can be tanh or ReLU.

The pooling layer downsamples the convolution output, extracts strong features and removes weak features, reduces the number of parameters, and prevents overfitting.

The fully connected layer performs regression classification on the features extracted from the previous layer. Through the weighted summation of the output of convolution and pooling layers and then through the response of activation function, the following formula is obtained:

$$x^l = f(w^l x^{l-1} + b^l), \quad (2)$$

where x^l is the network weight coefficient.

2.2. BiGRU. The BiGRU is a new type of the bidirectional recurrent neural network model. The recurrent neural network (RNN) effectively solves the problem that there is no operational connection between inputs in the hidden layer of traditional neural networks and can be used for the prediction of time series data and text semantic data. Its structural diagram is shown in Figure 2. However, the RNN uses a back propagation algorithm. When learning a long time sequence, the RNN may have gradient disappearance and gradient explosion problems, and thus, it is unable to grasp the nonlinear relationship of complex time series data with a long time span.

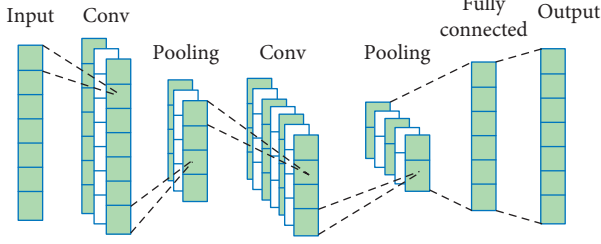


FIGURE 1: Structural diagram of the CNN.

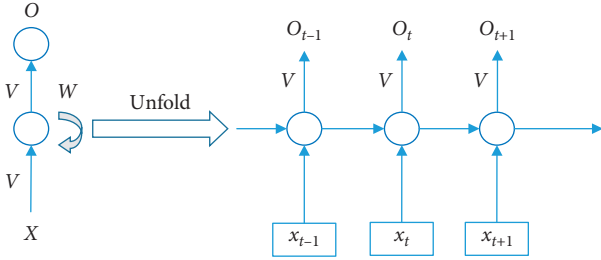


FIGURE 2: Structural diagram of an RNN unit.

When the recurrent neural network processes time series data, the transmission of its state is from front to back. However, in most complex time series data, the output at the current moment is not only related to the previous state but also related to the subsequent state. Schuster [22] proposed the bidirectional RNN (BiRNN) to solve the above problems. The basic idea is that each training sequence is completed by two forward and backward recurrent neural networks, and these two RNNs are connected to the same output layer. The output layer contains the complete information of the past and the future of each point in the input sequence, forming a double-loop network structure. The BiRNN brings a certain improvement compared with the ordinary RNN [23].

The bidirectional gated recurrent unit is a BiRNN based on the gated recurrent unit. In the BiGRU, two GRU inputs in opposite directions are provided at the same time at each time t . GRUs in the two directions are not directly connected, and the output is jointly determined by two unidirectional GRUs. The BiGRU model has good prediction performance in nonlinear time series data [24, 25]. The structure of the BiGRU is shown in Figure 3.

In the forward layer, we calculate the forward direction from time 1 to time t and obtain and save the output h_t of the forward hidden layer at each time. In the backward layer, the reverse calculation is performed from time t to time $t-1$, and the output h'_t of the backward hidden layer at each time is obtained and saved. Finally, the final output o_t is obtained by combining the output results of the forward layer and backward layer at each time. The mathematical expressions are presented as follows.

- (1) In the forward layer, the output h_t from front to back is

$$h_t = f(w_1 x_t + w_2 h_{t-1}). \quad (3)$$

- (2) In the backward layer, the output h'_t from back to front is

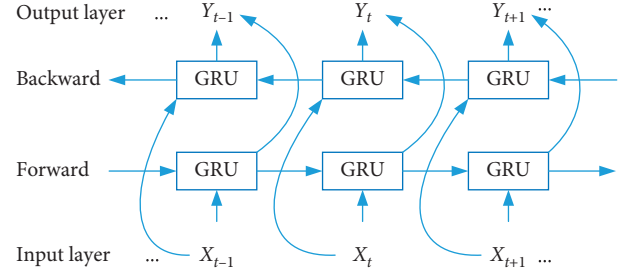


FIGURE 3: Structural diagram of the BiGRU.

$$h'_t = f(w_3 x_t + w_5 h'_{t-1}). \quad (4)$$

- (3) Combining the forward layer and the backward layer, the final output of the BiGRU is

$$o_t = g(w_4 h_t + w_6 h'_t), \quad (5)$$

where w_1 , w_2 , w_3 , w_4 , w_5 , and w_6 , respectively, represent the weights corresponding to the forward and reverse hidden states.

2.3. Model Building. Deep learning-based models have the ability to automatically extract features. CNN models usually rely on the convolution kernel of the convolutional layer to extract features. However, the existence of the convolution kernel limits the long-term dependency problem of the CNN when processing time series data [26]. In this study, the introduction of the BiGRU can effectively solve this problem, and we can capture the dependencies before and after the time series. In view of the high feature dimension of amino acid time series, we first use the CNN to extract the features of amino acid time series. Then, the BiGRU is used to process the feature vector of the amino acid sequence. Finally, amino acid sequence identification is achieved through the fully connected network. The model structure is shown in Figure 4.

This model is composed of three modules: the CNN, the BiGRU, and the recognition network. The CNN consists of five convolutional layers and five average pooling layers. The two-way long- and short-term memory network is composed of three layers of BiGRUs, and each layer is composed of 512, 256, and 96 hidden units. The identification network consists of one discarding layer (the discarding rate parameter is set to 0.35), one fully connected layer, and one Softmax. Finally, an amino acid recognition model is obtained.

3. Experimental Equipment and Samples

3.1. Experimental Equipment and Sample Preparation. The experimental device uses a fiber-type terahertz time-domain spectroscopy system (THz-TDS) with a signal-to-noise ratio of up to 70 dB, as shown in Figure 5. The device consists of a femtosecond laser, an antenna that excites and detects terahertz radiation, a delay line, and a lock-in amplifier. The center wavelength of the ultrashort pulse fiber

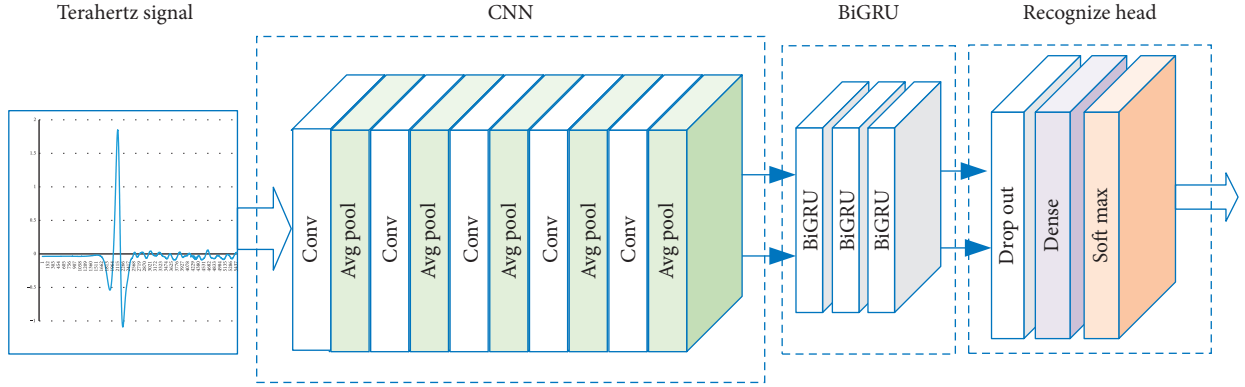


FIGURE 4: Structural diagram of amino acid terahertz spectrum recognition model based on the CNN and BiGRU.

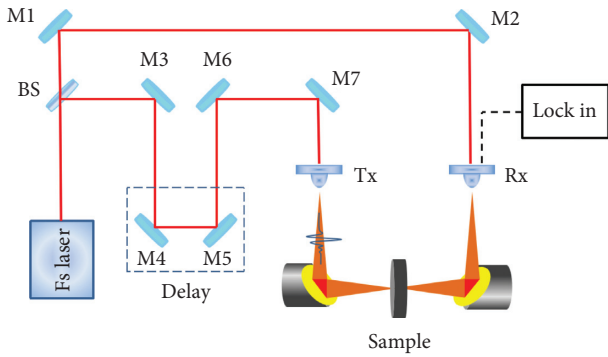


FIGURE 5: Schematic diagram of the terahertz time-domain spectroscopy system.

laser pulse is 1560 nm, and the repetition frequency is 100 MHz. In order to obtain higher resolution, the measurement time is 53 ps.

The three amino acid samples of glutamic acid, glutamine, and asparagine were provided by Shanghai Aladdin Reagent Company. The samples were baked for 24 h at 50°C and then ground with agate. The particle size was less than 80 μm to reduce the scattering effect. Samples were then mixed with high-density polyethylene powder in different proportions (30 different proportions), and tablets with a diameter of 13 mm were put under a pressure of 20 MPa [20]. The weight of each sample is 120 mg, and the thickness is about 1.2 mm. Each sample was measured at different times to obtain 10 terahertz spectra.

3.2. Evaluation Index. This study uses three evaluation indicators, that is, accuracy, recall, and precision, to evaluate the performance of the model. Accuracy is the proportion of amino acids correctly classified in the total test sample. Accuracy is the most intuitive way to evaluate the performance of the model.

Recall is the true-positive rate (TPR), that is, recall is the proportion of real positive samples to all positive samples that are currently classified into the positive sample category:

$$R = \frac{tp}{tp + fn}. \quad (6)$$

Precision is a measure of accuracy, which represents the proportion of examples that are divided into positive examples:

$$P = \frac{tp}{tp + fp}. \quad (7)$$

F-score is the weighted harmonic average of precision and recall:

$$F = \frac{(a^2 + 1)P * R}{a^2(P + R)}, \quad (8)$$

when $\alpha = 1$, then the F-score is F1:

$$F1 = \frac{2PR}{P + R}. \quad (9)$$

Area under curve is defined as the area under the receiver operating characteristic (ROC) curve and surrounded by the coordinate axis. For the ROC, the abscissa is the false-positive rate (FPR), and the ordinate is the TPR; therefore, when the TPR is larger and the FPR is smaller, the classification result is better.

4. Results and Discussion

In order to verify the effectiveness of the CNN-BiGRU model proposed in this article, we use the BiGRU [27], PCA-SVM [28], PCA-LSTM [29], and CNN-LSTM [30] for comparison. In the experiment, the training and validation sets totaled 1000, including 80% of the training set, 10% of the validation set, and 10% of the test set. The accuracy rate, recall rate, precision rate, and F1-score for each model are given in Table 1.

The ROCs of each model are shown in Figures 6–9.

The nonlinear fitting ability of the traditional machine learning method is very limited, and it may not be able to extract high-level and high-resolution features accurately. On the contrary, it will omit important information in denoising and feature extraction. At the same time, the effect of the SVM in classification is general, and thus, the PCA-SVM model is the worst in all indicators.

LSTM is specially designed for time series, but LSTM alone does not achieve good results because it overfits the

TABLE 1: Summary of evaluation indices of each model.

Model/index	Accuracy rate	Recall rate (%)	Precision (%)	F1-score (%)
PCA-SVM	68.35%	68.35	68.35	56.08
BiGRU	91.56%	91.56	91.56	97.82
PCA-LSTM	97.89%	97.89	97.89	97.82
CNN-BiGRU	99.16%	99.16	99.16	99.17

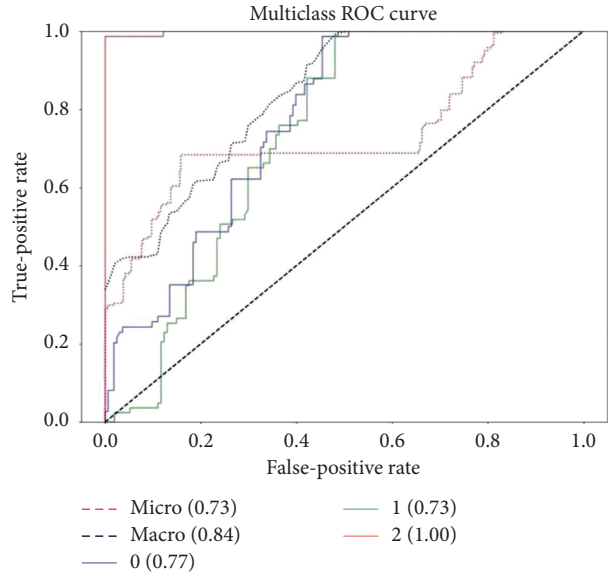


FIGURE 6: PCA-SVM ROC.

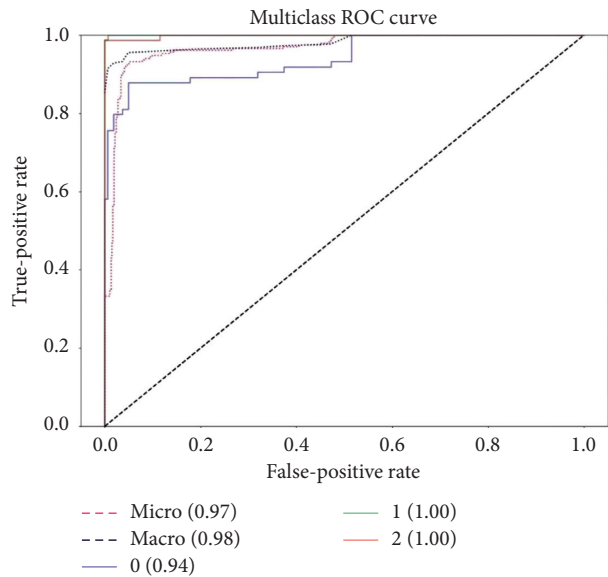


FIGURE 7: PCA-LSTM ROC.

sequence characteristics of amino acids. The main reason why PCA-LSTM is better than LSTM is that PCA compresses features and eliminates most redundant data.

The CNN-BiGRU model was used to input the amino acids after a simple pretreatment. The morphological features of basic acids were extracted by the CNN, and then, the temporal features of amino acids were extracted by the

BiGRU. The two features were combined to mine the hidden deep information in amino acids. The average accuracy of classification test results was 99.16%. It can be seen from table that our model achieved the best results in all indicators. The main reason for this is that the CNN can provide more features, and the BiGRU can consider the relationship between features.

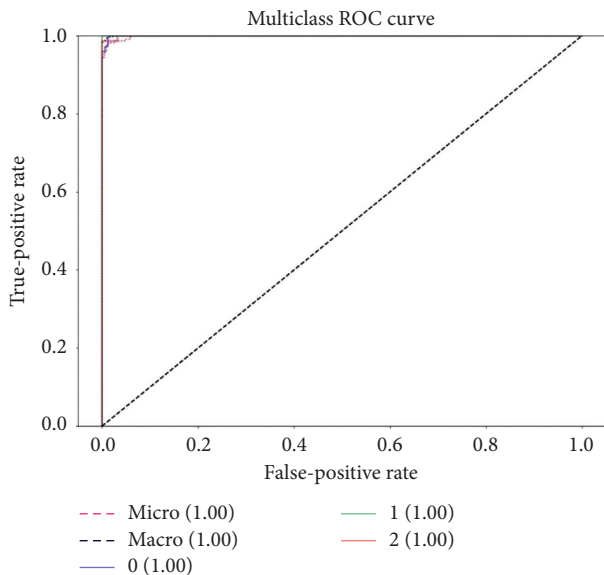


FIGURE 8: BiGRU ROC.

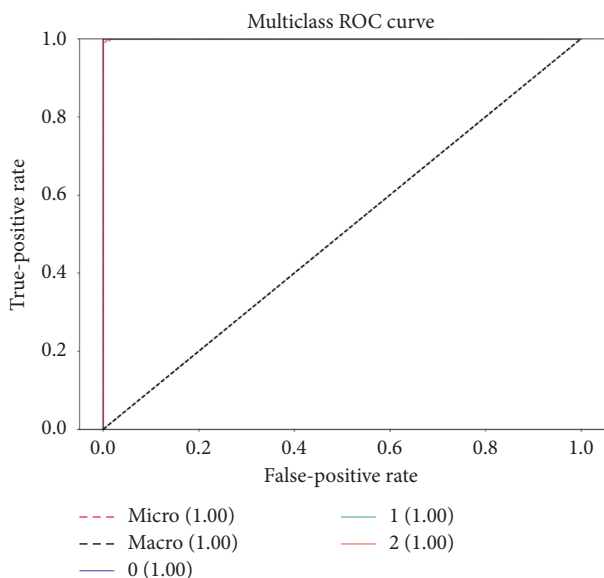


FIGURE 9: CNN-BiGRU ROC.

5. Conclusion

Because it is difficult to conduct a large number of experiments in the experimental environment, the number of samples of various amino acids is small in the present study, which has a certain impact on the experimental results. However, the relative effect of the tested models is not affected. In future, we will conduct more experiments and provide samples of more types of amino acids.

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

This research was supported by the National Key R&D Program of China (2018YFC0809200).

References

- [1] P. H. Siegel, "Terahertz technology," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 910–928, 2002.
- [2] Y. Xiang, Z. Xiang, Y. Ke et al., "Biomedical applications of terahertz spectroscopy and imaging," *Trends in Biotechnology*, vol. 34, no. 10, pp. 810–824, 2016.
- [3] K. Serita, E. Matsuda, K. Okada, H. Murakami, I. Kawayama, and M. Tonouchi, "Invited Article: terahertz microfluidic chips sensitivity-enhanced with a few arrays of meta-atoms," *Appl Photonics*, vol. 3, no. 5, Article ID 051603, 2018.
- [4] W. Zhang, E. R. Brown, M. Rahman, and M. L. Norton, "Observation of terahertz absorption signatures in microliter DNA solutions," *Applied Physics Letters*, vol. 102, no. 2, pp. 219–229, 2013.
- [5] E. R. Brown, E. A. Mendoza, D. Xia, and S. R. J. Brueck, "Narrow THz spectral signatures through an RNA solution in nanofluidic channels," *IEEE Sensors Journal*, vol. 10, no. 3, pp. 755–759, 2010.
- [6] M. R. Kutteruf, C. M. Brown, L. K. Iwaki, M. B. Campbell, T. M. Korter, and E. J. Heilweil, "Terahertz spectroscopy of short-chain polypeptides," *Chemical Physics Letters*, vol. 375, no. 3–4, pp. 337–343, 2003.
- [7] L. Xie, Y. Yao, Y. Ying et al., "The application of terahertz spectroscopy to protein detection: a review," *Applied Spectroscopy Reviews*, vol. 49, no. 6, pp. 448–461, 2014.
- [8] J. Neu, E. A. Stone, J. A. Spies et al., "Terahertz spectroscopy of tetrameric peptides," *The Journal of Physical Chemistry Letters*, vol. 10, no. 10, pp. 2624–2628, 2019.
- [9] S. Lu, X. Zhang, Z. Zhang, Y. Yang, and Y. Xiang, "Quantitative measurements of binary amino acids mixtures in yellow foxtail millet by terahertz time domain spectroscopy," *Food Chemistry*, vol. 211, pp. 494–501, 2016.
- [10] J. El Haddad, B. Bousquet, L. Canioni, and P. Mounaix, "Review in terahertz spectral analysis," *TrAC Trends in Analytical Chemistry*, vol. 44, pp. 98–105, 2013.
- [11] J. El Haddad, F. de Miollis, J. Bou Sleiman, L. Canioni, P. Mounaix, and B. Bousquet, "Chemometrics applied to quantitative analysis of ternary mixtures by terahertz spectroscopy," *Analytical Chemistry*, vol. 86, no. 10, pp. 4927–4933, 2014.
- [12] J. B. Sleiman, B. Bousquet, N. Palka, and P. Mounaix, "Quantitative analysis of hexahydro-1,3,5-trinitro-1,3,5, triazine/pentaerythritol tetranitrate (RDX-PETN) mixtures by terahertz time domain spectroscopy," *Applied Spectroscopy*, vol. 69, no. 12, pp. 1464–1471, 2015.
- [13] Y. Ueno, R. Rungsawang, I. Tomita, and K. Ajito, "Quantitative measurements of amino acids by terahertz time-domain transmission spectroscopy," *Analytical Chemistry*, vol. 78, no. 15, pp. 5424–5428, 2006.
- [14] L. Yuan, L. Bin, and L. Huan, "Analysis of fluoroquinolones antibiotic residue in feed matrices using terahertz spectroscopy," *Applied Optics*, vol. 57, no. 3, p. 544, 2018.

- [15] Y. Peng, C. Shi, M. Xu et al., "Qualitative and quantitative identification of components in mixture by terahertz spectroscopy," *IEEE Transactions on Terahertz Science and Technology*, vol. 8, no. 6, pp. 696–701, 2018.
- [16] W. Liu, C. Liu, X. Hu, J. Yang, and L. Zheng, "Application of terahertz spectroscopy imaging for discrimination of transgenic rice seeds with chemometrics," *Food Chemistry*, vol. 210, no. 1, pp. 415–421, 2016.
- [17] Y. Xu, G. Yang, J. Luo, and J. He, "An electronic component recognition algorithm based on deep learning with a faster SqueezeNet," *Mathematical Problems in Engineering*, vol. 2020, Article ID 2940286, 11 pages, 2020.
- [18] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [19] K. Xu, J. Ba, R. Kiros et al., "Show, attend and tell: neural image caption generation with visual attention," *Proceedings of International Conference on Machine Learning*, vol. 37, no. 7, pp. 2048–2057, 2015.
- [20] Y. Xu, W. Yan, G. Yang, J. Luo, T. Li, and J. He, "CenterFace: joint face detection and alignment using face as point," *Scientific Programming*, vol. 2020, Article ID 7845384, 8 pages, 2020.
- [21] K. Ping-Huan and H. Chiou-Jye, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, pp. 213–226, 2018.
- [22] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions On Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [23] D. She and M. Jia, "A bigru method for remaining useful life prediction of machinery," *Measurement*, vol. 167, no. 1, Article ID 108277, 2020.
- [24] Q. Zhu, F. Zhang, S. Liu et al., "A hybrid VMD-bigru model for rubber futures time series forecasting," *Applied Soft Computing*, vol. 84, no. 1, Article ID 105739, 2019.
- [25] Q. Zhu, F. Zhang, S. Liu, and Y. Li, "An anticrime information support system design: application of K-Means-VMD-Bigru in the city of chicago," *Information & Management*, vol. 17, no. 11, Article ID 103247, 2019.
- [26] D. U. Yongping, X. Zhao, and B. Pei, "Short text sentiment classification based on CNN-LSTM model," *Journal of Beijing University of Technology*, vol. 45, no. 7, pp. 662–670, 2019.
- [27] A. Baccouche, B. Garcia-Zapirain, C. Castillo Olea, and A. Elmaghraby, "Ensemble deep learning models for heart disease classification: a case study from Mexico," *Information*, vol. 11, no. 4, p. 207, 2020.
- [28] L. I. Yun-Fei, Y. Q. Huang, and G. L. Jiang, "Short-term load forecasting based on PCA-SVM," *Proceedings of the Chinese Society of Universities for Electric Power System and Automation*, vol. 19, no. 5, pp. 66–70, 2007.
- [29] M. Zimmermann, M. Mehdipour Ghazi, H. K. Ekenel, and J. P. Thiran, "Combining multiple views for visual speech recognition," in *Proceedings of International Conference on Auditory-Visual Speech Processing (AVSP)*, Stockholm, Sweden, August 2017.
- [30] T. N. Nguyen, D. Q. Tran, T. N. Nguyen, and H. Q. Nguyen, "A CNN-LSTM architecture for detection of intracranial hemorrhage on CT scans," arXiv preprint arXiv:2005.10992, 2020.

Research Article

A Medical Image Classification Model Based on Adversarial Lesion Enhancement

Bing Zhang ¹ and Xu Hu²

¹*School of Information Engineering, Tianjin University of Commerce, Tianjin, China*

²*School of Economics and Management, Tsinghua University, Beijing, China*

Correspondence should be addressed to Bing Zhang; zhangbing@tju.edu.cn

Received 21 April 2021; Accepted 17 May 2021; Published 28 May 2021

Academic Editor: Wenbing Zhao

Copyright © 2021 Bing Zhang and Xu Hu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of Artificial Intelligence, the auxiliary diagnosis model based on deep learning can assist doctors to a certain extent. However, the latent information in medical images, such as lesion features, is ignored in most of the traditional methods. The extraction of this information is regarded as a learning task within the network in some recent researches, but it requires a large amount of fine-labeled data, which is undoubtedly expensive. In response to the problem above, this paper proposes an Adversarial Lesion Enhancement Neural Network for Medical Image Classification (ALENN), which is used to locate and enhance the lesion information in medical images only under weakly annotated data so as to improve the accuracy of the auxiliary diagnosis model. This method is a two-stage framework, including a structure-based lesion adversarial inpainting module and a lesion enhancement classification module. The first stage is used to repair the lesion area in the images while the second stage is used to locate the lesion area and use the lesion enhanced data during modeling process. In the end, we verified the effectiveness of our method on the MURA dataset, a musculoskeletal X-ray dataset released by Stanford University. Experimental results show that our method can not only locate the lesion area but also improve the effectiveness of the auxiliary diagnosis model.

1. Introduction

In December 2012, a study [1] showed that musculoskeletal diseases, such as arthritis and back pain, are the second-leading cause of disability as well as the fourth-leading factor to the health of the world population, affecting more than 1.7 billion people worldwide. According to data from the World Health Organization [2], there are more than 150 diseases caused by the musculoskeletal (exercise) system. Although diseases generated by different reasons (for example, exercise, genetics, or poor lifestyle) have little difference in visual perception, there is a big difference among the disease types, so it requires different diagnosis and treatment options. Therefore, the detection of musculoskeletal abnormalities appears particularly important. The abnormalities of musculoskeletal are mainly reflected in the basic diseases of bones, joints and soft tissues [3, 4]. Among them, basic bone diseases include osteoporosis, osteomalacia, bone destruction, osteosclerosis, periosteal proliferation, chondral

calcification, osteonecrosis, bone deformity, etc. Basic joint diseases include swelling of joint, destruction of joint, degeneration of joint, ankylosis of joint, dislocation of joint, etc. Basic soft tissue diseases include soft tissue swelling, soft tissue mass, myophagism, etc.

Bone is the tissue with the highest density in the human structure, which has a clear contrast with surrounding tissues. Meanwhile, there also exists an obvious contrast between the cortex and cancellous bone of the bone itself so that the conventional X-ray examination can be used for general bone diseases diagnosis. In addition, due to the advancement of imaging technology and the upgrading of imaging equipment, various hospitals have produced a large amount of medical imaging data, and these precious data are helpful for many researches. Therefore, it is of great research significance to use computer-aided diagnosis technology to quickly and accurately classify musculoskeletal diseases based on a large number of existing medical images.

Many machine learning methods have been applied to medical image data classification tasks, including K-Means clustering [5, 6], decision tree [7], support vector machine [8], and random forest [9]. However, the number of features extracted by traditional machine algorithms is limited, and only artificially preset features can be classified. Therefore, traditional machine learning algorithms perform poorly on medical image classification. With the continuous development of deep learning technology, deep learning algorithms, including CNN (convolutional neural networks) [10] and GAN (generative adversarial networks) [11], are widely used in classification tasks. Through the fine-grained annotation from a large number of professional doctors, the CNN model can automatically extract features and can extract more feature information through a deeper and wider network architecture. But the difficulty of medical image annotation makes it more expensive to improve CNN model performance. In addition, most of the existing CNN models use benign and malignant labeling information as the last layer of supervision information on the network, and they cannot dig out more hidden features through coarse-grained benign and malignant labeling to reuse and merge key features.

Based on the problems above, this paper proposes an Adversarial Lesion Enhancement Neural Network for Medical Image Classification (ALENN), which automatically recognizes the lesion area in the image only through the supervision of category annotation, and enhances and optimizes the prediction accuracy of the auxiliary diagnosis model. This method is a two-stage model. The first stage is a structure-based lesion adversarial inpainting module, which is used to repair the lesion area in the image; the second stage is a lesion enhancement classification module, which is used to identify the lesion area and apply the data after the enhanced lesion to assist the modeling process of the diagnostic model. The core of the first stage is structural information, which represents the relatively fixed semantics of human body structure in medical images. We believe that better restoration results can be obtained by splitting the image restoration process into structural semantic restoration and texture detail restoration. The core of the second stage is sliding window. Through the sliding of the occluded area, the most significant abnormal area in the image can be found. Finally, we verified the effectiveness of this method on the MURA dataset [12], a musculoskeletal X-ray film dataset released by Stanford University.

The related works of traditional and latest auxiliary diagnosis as well as the related researches on the MURA dataset are introduced in the second section of this paper. The method proposed in this paper is introduced in the third section. The experimental results and analysis are shown in the fourth section. The summary and prospects of the work are given in the fifth section.

2. Related Work

CNN-based medical image analysis method has shown excellent performance in many challenging tasks (disease classification [13], lesion detection [14], fine-grained lesion

segmentation [15]), among which it has deepest research and is most widely applied in disease classification. X-rays examination is one of the most common radiology examinations in the clinical diagnosis of chest diseases. How to combine a large number of existing medical images with the rich clinical experience of professionally trained radiologists is of great significance for the diagnosis of chest diseases. Wang et al. [16] proposed a novel text-image embedding network (TieNet) to extract images and corresponding text representations. It employs an end-to-end trainable CNN-RNN architecture embedded with a multi-level attention model to highlight important image regions and their corresponding text words, and then use image features and text embeddings extracted from related reports to classify the chest X-rays. Coronary angiography is the gold standard for computer-aided diagnosis (CAD), so it is essential to describe in detail the position and the degree of stenosis through coronary angiography for classifying the severity of CAD. Wang et al. [17] used recursive capsule network (RCN) to extract the semantic relationship between clinically named entities in the coronary angiography text, so as to automatically find out the maximum stenosis degree of each lumen, and finally inferred the coronary artery severity according to the improved Gensini method.

The MURA dataset is the largest public musculoskeletal image dataset available currently. Many scholars have conducted numerous experimental studies on this dataset, including the use of traditional machine learning methods and deep learning algorithms. Among them, Pawan et al. [18] used support vector machine (SVM), linear SVM, logistic regression, and decision tree algorithms to detect musculoskeletal anomalies, and introduced a gray-level co-occurrence matrix (GLCM) to preprocess the original musculoskeletal images, comparing them on five evaluation indicators of sensitivity, specificity, precision, accuracy, and F1-score in the end. Pranav et al. [19] designed a 169-layer baseline model DenseNet to solve the musculoskeletal abnormality detection. On the MURA data and above, DenseNet's performance is lower than the worst radiologist in 5 of the 7 basic studies, and the performance of the overall model is also lower than the best radiologist. Subsequently, Dennis et al. [2] proposed the idea of ensemble learning to integrate the well-trained classification models DenseNet201, MobileNet and NASNet-Mobile, and designed the ensemble200 model. Finally, its Cohen Kappa score is 0.66, which is lower than the DenseNet169-layer model. But the F1-score of ensemble200 is better than that of DenseNet model, and the Cohen Kappa score variance of different parts is even lower. Due to the sensitivity of CNN to extract features, when the image undergoes transformations, such as rotation and misalignment, the performance of CNN recognition will be greatly reduced. To solve this problem, SAIF et al. [20] introduced the capsule network architecture. Through its powerful dynamics routing mechanism, an output vector containing many feature information (including spatial direction, vector amplitude) can be obtained to achieve abnormal detection of musculoskeletal X-ray films, and this method only applies to a small number of network layers to use 169 densely connected CNNs Kappa

coefficient of the model. Although the CNN model has shown good performance on many datasets, it performs poorly on some subsets. Luke et al. analyzed that the reason is that the model cannot fully describe the complete changes in the dataset, and he defined this problem as hidden layering. Therefore, Luke et al. [21] evaluated several possible techniques for measuring the effect of hidden stratification and found that hidden stratification can occur in unrecognized low prevalence, low label quality, subtle distinguishing features, and it can lead to more than 20% relative performance difference of the models.

3. Method

Due to the inadequate utilization of lesion information by current deep learning models based on medical image datasets (e.g., MURA) and the inexplicability of deep learning itself, there is still much room for the reliability and credibility of auxiliary diagnosis models to improvement. Aiming at the problems above, this paper proposes an Adversarial Lesion Enhancement Neural Network for Medical Image Classification (ALENN). The method includes two main modules: a structure-based lesion adversarial inpainting module and a classification module based on lesion information fusion. The overall two-stage structure diagram is shown in Figure 1. The LE represents the lesion enhancement.

3.1. Structure-Based Lesion Adversarial Inpainting Module. The X-ray image of the elbow reflects the density difference among the different tissues of the elbow, and the density distribution in the negative data is obviously different from that in the positive data, as shown in Figure 2. The difference of the distribution is reflected in the lesion area in the positive data. This paper assumes that, in the positive data, the tissue density distribution outside the lesion area is similar to the negative data. Therefore, this distribution difference can be fully utilized to find the location of the lesion in the image, thereby providing additional information for the fusion of the lesion area. If there is a generator that only uses negative data for training, that is, it fits the distribution of negative data under ideal conditions, it can theoretically restore the lesion area in the positive data to the data distribution of normal tissue. In other words, the restored image can be regarded as a piece of pseudo-negative data with a distribution similar to the negative data. Hence, this module is based on the idea of the generative adversarial network to generate positive data into negative data, and regard the relative error between the data as important information for subsequent enhancement of the lesion area.

At the same time, in order to better restore the original semantic information of the image, considering that medical images often contain relatively fixed human structural features-bones and muscles, this module extracts the structural information of the data so that the model pays more attention to the relatively fixed structure, not limited to susceptible texture information. Consequently, inspired by the research [22], this paper adds the models generation and

learning process of structural information on the basis of the generative adversarial network.

3.1.1. Structure Information. For the extraction of structure information, we assume that the image is composed of structure and texture and uses the relative total variation (RTV) in [23] to distinguish the structure and texture in the image. The input image is defined as I , the pixel index in the image is defined as p , and the structure in the image is defined as S . Consequently, the process of using a secondary penalty to strengthen the structural similarity between input and output in the TV-L2-based model can be expressed as

$$\arg \min_s \sum_p \left\{ \frac{1}{\lambda} (S_p - I_p)^2 + |(\nabla S)_p| \right\}, \quad (1)$$

where ∇ represents the first-order difference operation, $(S_p - I_p)^2$ is used to maintain the similarity between the structure and the input image, and $|(\nabla S)_p|$ can be divided into two directions: x direction and y direction:

$$\sum_p |(\nabla S)_p| = \sum_p \left\{ |(\partial_x S)_p| + |(\partial_y S)_p| \right\}. \quad (2)$$

However, the authors of [23] found that the total variation regularizer has limited ability to distinguish between strong structural edges and texture. Therefore, in order not to target a certain type of texture structure image, the overall task can be rewritten as

$$\arg \min_s \sum_p (S_p - I_p)^2 + \lambda \cdot \left(\frac{G_x(p)}{N_x(p) + \varepsilon} + \frac{G_y(p)}{N_y(p) + \varepsilon} \right), \quad (3)$$

where λ is the weight and ε is a small positive number, which should avoid being divided by zero. $G(p)$ represents a general pixel-wise windowed total variation measure and $N(p)$ represents a novel windowed inherent variation, which can be expressed as

$$\begin{aligned} G_x(p) &= \sum_{q \in R(p)} g_{p,q} \cdot |(\partial_x S)_q|, \\ G_y(p) &= \sum_{q \in R(p)} g_{p,q} \cdot |(\partial_y S)_q|, \\ N_x(p) &= \left| \sum_{q \in R(p)} g_{p,q} \cdot (\partial_x S)_q \right|, \\ N_y(p) &= \left| \sum_{q \in R(p)} g_{p,q} \cdot (\partial_y S)_q \right|, \end{aligned} \quad (4)$$

where q is the index of all pixels in a square area centered on point p and $g_{p,q}$ is a weighting function defined according to spatial affinity. However, due to the fact that objective function is non-convex and its solution cannot be obtained directly, a numerically stable approximation of the solution needs to be obtained by decomposing the nonlinear part and the quadratic part.

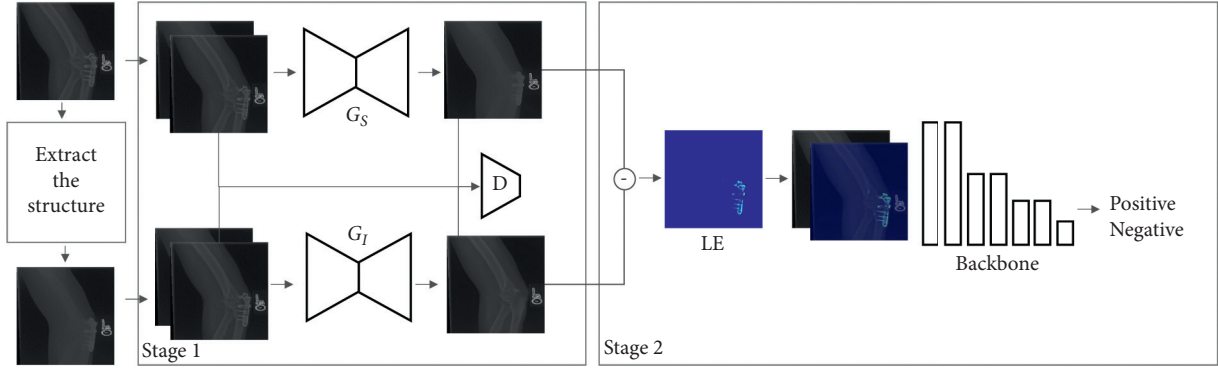


FIGURE 1: The overall structure of the two-stage ALENN module.

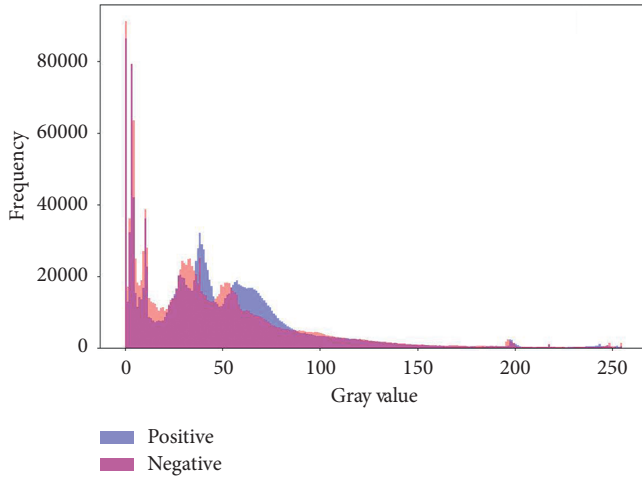


FIGURE 2: Data distribution in positive and negative samples.

3.1.2. GAN. In the architecture of this paper, the role of the generative adversarial network is to provide the lesion area in the elbow X-ray image for the final classification model. We additionally assume the MURA data: the distribution difference between the positive data and the negative data in the image results from the lesion. If the lesion area in the positive image is completely occluded, the part of the image that is not occluded at this time is similar in distribution to the negative data. The structure information obtained in Section 3.1.1 plays a significant role in GAN, which forces the generator to ignore the local interference caused by texture information, and then reconstruct the physiological structure information inside the image. And it fills and repairs the detailed texture on the basis of reconstructed structure information. More specifically, the image structure is used as the supervision information firstly, and the GAN is trained through the negative image repair task occluded by the random mask, so that the generator first fits the structure distribution of the bone and muscle in the negative data. On this basis, the original image is used as supervisory information to train GAN, so that the generator gradually fills in the detailed texture information on the basis of the structure. In the process of training the generator, the discriminator is also trained at the same time to judge the authenticity of the generated image and fight against the generator. Finally, the

positive data is used as the test data, and the lesion features in the image are repaired on the premise of preserving normal tissues, and then restored to pseudo-negative data.

This section follows the definitions of I and S in Section 3.1.1 and adds other definitions: generator G , discriminator D , and binarized matrix mask M . In addition, the image I is further expressed as $I = I_p \cup I_n$, where I_p represents the positive sample of the lesion in the image, and I_n represents a negative sample. Similarly, $S = S_p \cup S_n$. Based on this, the output of the GAN used to repair the image structure information in the first step can be expressed as

$$\hat{S} = G_{I \rightarrow 5}(E(I_n, M), E(S_n, M)), \quad (5)$$

where $E(x, M) = x \odot M$ represents the occlusion operation; that is, the area with the value of 0 in M is occluded in the corresponding input data x . \odot denotes element-wise product. At the same time, the discriminator needs to judge the authenticity of the image and conduct adversarial training with the generator, and its output can be expressed as $T_S = D_S(S)$ and $\hat{T}_S = D_S(\hat{S})$.

In addition to training the generator to make the overall model capable to reconstruct structural information, it is also necessary to train the generator, so that the overall model has the ability to gradually repair texture information through structural information. And because this paper assumes “image = structure + texture,” after the texture information is supplemented and perfected on the basis of the structure information, it can be regarded as the overall restoration of the image. Therefore, the output of the generator $G_{S \rightarrow I}$ can be expressed as

$$\hat{I} = G_{s \rightarrow f}(E(I_n, M), \hat{S}). \quad (6)$$

Same as the repair structure stage, the discriminator still needs to predict the authenticity of the generator output image, and the output at this time is expressed as $T_I = D_I(I)$ and $\hat{T}_I = D_I(\hat{I})$.

3.1.3. Loss Function. The GAN in this paper requires two-step training. In other words, the training model firstly repairs the structure image with complex texture semantics removed. After the repair process is smooth, the training model will repair the detailed texture information based on

the structure. In the first training step, in order to fit the true distribution of the image structure S_n , we apply the idea of confrontation to the framework of the generative model. The

adversarial loss of the generator and the discriminator can be written as

$$L_{G,I \rightarrow \hat{S}} = \|S - \hat{S}\|_1, \quad (7)$$

$$\begin{aligned} L_{D,S} &= -\sum p(S) \log q(\hat{S}) = \mathbb{E}[\log(T_S)] + \mathbb{E}[\log(1 - \hat{T}_S)] \\ &= \mathbb{E}[\log(D_S(S))] + \mathbb{E}\left[\log\left(1 - D_S\left(G_{I \rightarrow \hat{S}}(E(I_n, M), E(S_n, M))\right)\right)\right]. \end{aligned} \quad (8)$$

Similar to the first training process, the loss function for the repair process of training GAN on texture details is as follows:

$$L_{G,\hat{S} \rightarrow \hat{I}} = \|I - \hat{I}\|_1, \quad (9)$$

$$\begin{aligned} L_{D,I} &= -\sum p(I) \log q(\hat{I}) = \mathbb{E}[\log(T_I)] + \mathbb{E}[\log(1 - \hat{T}_I)] \\ &= \mathbb{E}[\log(D_I(S))] + \mathbb{E}\left[\log\left(1 - D_I\left(G_{\hat{S} \rightarrow \hat{I}}(E(I_n, M), \hat{S})\right)\right)\right]. \end{aligned} \quad (10)$$

The two stages of adversarial losses are combined through hyper-parameters α and β , respectively, and go through the final optimization process:

$$L_S = \alpha L_{G,I \rightarrow \hat{S}} + (1 - \alpha) L_{D,S}, \quad (11)$$

$$L_I = \beta L_{G,\hat{S} \rightarrow \hat{I}} + (1 - \beta) L_{D,I}. \quad (12)$$

We set $\alpha = 0.8$ and $\beta = 0.85$ in this paper.

3.2. Lesion Fusion Classification Module. This section will introduce the second stage of ALENN, which is the medical image-assisted diagnosis module based on lesion information enhancement. If there is an auxiliary diagnosis model based only on the original data modeling, this model can well complete the auxiliary diagnosis task. However, due to the inexplicability of deep learning itself, the model can easily fall into a local optimal value. In addition, there also exist certain limitations on the convergence speed of the model, which depends on the model's initialization parameters and optimization strategies. Hence, the relative error between the positive data and the negative data and the similarity between the positive data are used to artificially amplify this distribution difference at this stage. More specifically, after the restoration of the positive data in the first stage, the corresponding pseudo-negative data is obtained, while the real negative data can still be regarded as negative data after being repaired. Based on this, we can enhance the lesion area in the data and then use it as the

modeling data for the auxiliary diagnosis model, making the model sensitive to differences in distribution.

The lesion fusion assisted diagnosis proposed in this paper is an overall framework, and the optimization of the specific network structure is beyond the scope of this paper. In other words, the method proposed in this paper is applicable to most classification networks. Therefore, we select the most basic model VGG-19 as the backbone of this framework. Next, we will gradually introduce the overall process of lesion information fusion in the second stage. In order to facilitate the presentation, we define the subsequent variables. Input image I , where W and H represent the width and height of the image, respectively. Set the sliding window set $D = \cup_i \text{window}_i$ for each input sample; the window size of the sliding window is w and h , which is a hyper-parameter set according to a priori. Similarly, the step size of the sliding window d is the hyper-parameter. Generally, we set $d < w$ so that each pixel in the input sample can get as many repair results as possible to reduce the impact of generation errors. From this, we can conclude that the number of sliding windows for an input sample is approximately $N = ((W - w) \times (H - h)) / d^2$. The lesion location algorithm based on sliding window is shown in Algorithm 1.

Consequently, under the supervision of category labels, we have completed the extraction of lesion information in medical images based on the difference of conditional probability distribution $P(I|\text{positive})$ and $P(I|\text{negative})$. Naturally, the artificial enhancement of this information makes the final auxiliary diagnosis model more sensitive to the difference in distribution between positive and negative

data. From another perspective, the enhancement of this lesion information can be seen as a priori attention to the lesion area to a certain extent. And this “prior” is not supervised, let alone man-made, but calculated from the difference in conditional probability distribution on image-level annotations.

4. The Experimental Results

4.1. Dataset and Training Setting. The MURA dataset is the largest public dataset of musculoskeletal image currently, jointly released by the Department of Computer Science, Medicine and Radiology of Stanford University. It contains a total of X training sets and X verification sets. The dataset selected 40,561 multi-view X-rays images of 12,173 patients from Stanford Hospital from 2001 to 2012 as sample data and was marked as normal or abnormal by professional radiologists. The result showed that 62% of the images were normal data while 38% were abnormal ones. The dataset consists of research types of finger, elbow, hand, humerus, forearm shoulder, and wrist. This paper only uses elbow data for experimental verification. And in order to show the experimental results fairly, the unified optimizer Adam is employed in this paper and the learning rate is initialized to 0.0001 with the batch size set to 16.

4.2. Evaluation Metric. In order to better evaluate the method proposed in this paper, we use peak signal-to-noise ratio (PSNR) coefficient and structural similarity (SSIM) to evaluate the performance of the repair module. The larger the value of the two indicators, the more similar the two images. And the accuracy, sensitivity, specificity, recall, F1-score, and Kappa score are applied to evaluate the performance of the classification module.

PSNR coefficient and SSIM are the most widely used objective measurement methods for evaluating repaired images. The calculation formula is as follows:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right), \quad (13)$$

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2\mu_y^2 + c_1)(\sigma_x^2\sigma_y^2 + c_2)}. \quad (14)$$

In this formula, μ_x is the average of x and μ_y is the average of y . σ_x^2 is the variance of x , σ_y^2 is the variance of y , and σ_{xy} is the variance of x and y . $c_1 = (k_1 L)^2$ and $c_2 = (k_2 L)^2$ are used to maintain stability. L is the dynamic range of pixel values. $k_1 = 0.01$ and $k_2 = 0.03$. The range of structural similarity is 0 to 1. When the two images are exactly the same, the value of SSIM is equal to 1.

In addition, if FP, FN, TP, and TN represent the false positive rate, false negative rate, true positive rate, and true negative rate, respectively, the calculation formulas for accuracy, recall, precision, F1-score, and Kappa score are as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (15)$$

$$\text{recall} = \frac{TP}{TP + FN}, \quad (16)$$

$$\text{precision} = \frac{TP}{TP + FP}, \quad (17)$$

$$F1 = 2 \cdot \frac{\text{precision} + \text{recall}}{\text{precision} + \text{recall}}, \quad (18)$$

$$\text{Kappa} = \frac{p_o - p_e}{1 - p_e}. \quad (19)$$

In this formula, p_o is the sum of the number of samples correctly classified in each category divided by the total number of samples, which is the overall classification accuracy. Assuming that the number of real samples in each category is a_1, a_2, \dots, a_n , respectively, and the predicted number of samples in each category is b_1, b_2, \dots, b_n , respectively, and the total number of samples is n , then $p_e = (a_1 \cdot b_1 + a_2 \cdot b_2 + \dots + a_n \cdot b_n) / (n \cdot n)$.

4.3. The Results of Lesion Enhancement. In order to verify the effectiveness of the first stage of the two-stage lesion enhancement classification method proposed in this paper, firstly it is necessary to qualitatively analyze the repair effect of the generated adversarial network. This section will perform visual analysis on the negative data and positive data of elbow data in MURA. Since the GAN in this paper is modeled based on negative data, in theory, GAN will approximate the distribution of negative data. On the contrary, when positive data is used as the input of GAN, GAN will repair the “abnormal” area in the positive data according to the negative data. The results of repair and visualization of negative and positive data are shown in Figure 3, in which the Difference indicates the difference between the original images and the restored ones and LE represents that the difference area is enhanced and displayed in the form of a heat map. The last column combines LE with the original images for more intuitive observation to the effectiveness of this method.

After qualitative analysis, we will quantitatively analyze the differences between the two data types, as shown in Table 1. It can be clearly seen that the repair effects of GAN are different for the two distributed images. However, this difference in distribution is directly reflected on the image and enhanced in this paper so as to provide clear guidance for auxiliary diagnosis of the two types of data in the next stage.

4.4. The Results of Classification. In order to prove the effectiveness and excellent adaptability of the method proposed in this paper, this section applies this method to each classification network. VGG-19, ResNet-50, DenseNet-121, and Inception-v3 are used to conduct related experiments and analyze the results. The experimental results are shown in Table 2.

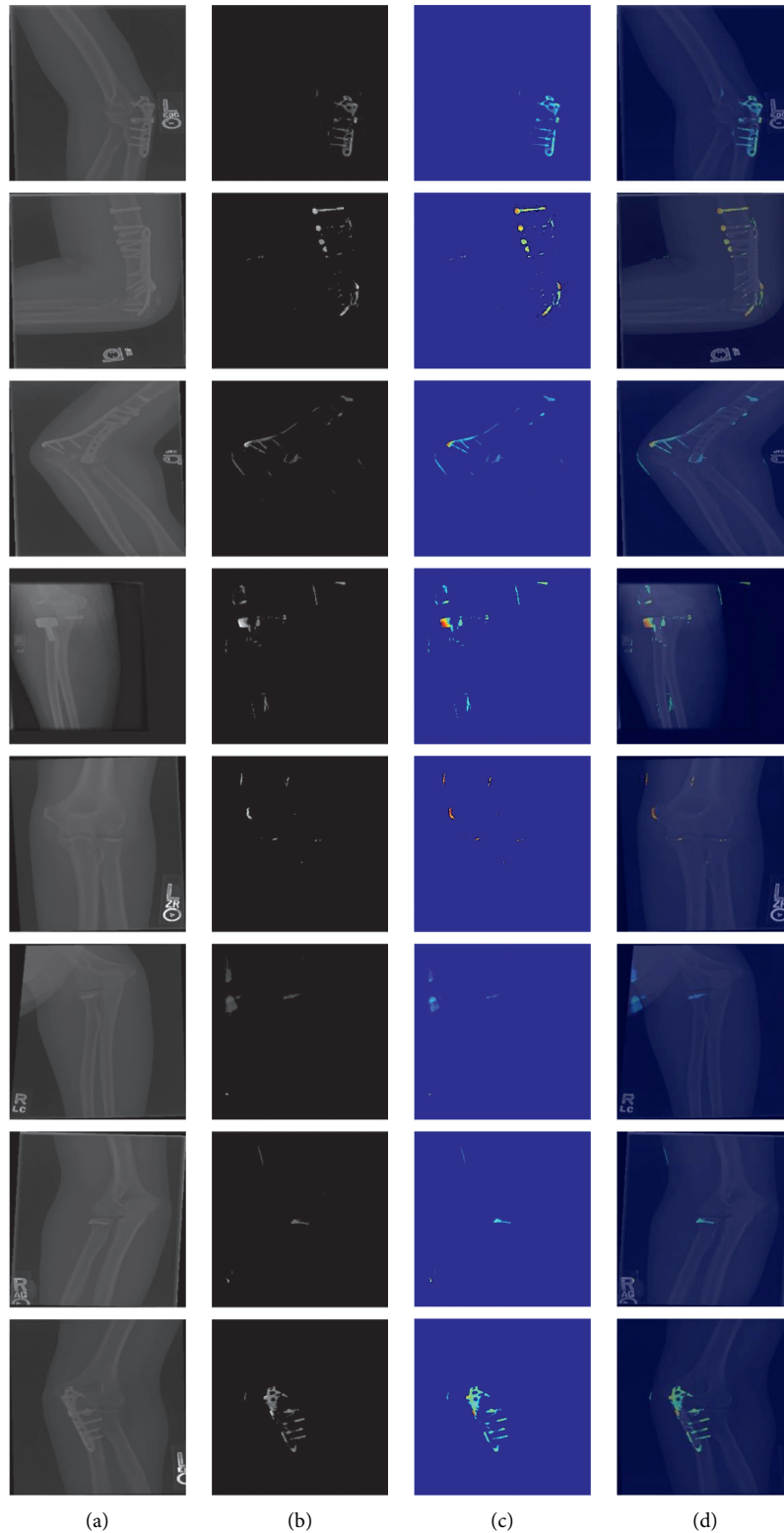


FIGURE 3: The results of repair and visualization of negative and positive data. (a) Original image. (b) Difference. (c) LE. (d) Color map.

It can be seen that the results are unsatisfactory compared to the method using only the classification network and the LE information. For each backbone, this method can increase the accuracy by about 1.67% on average. Moreover,

the method in this paper is the best one to lift VGG-19. The ROC curve of ResNet and VGG can explain this problem from another aspect, as shown in Figure 4. It is probably caused by the network structure of VGG itself. In other

TABLE 1: The differences between the two data types.

	PSNR	SSIM
Negative	34.7566	0.9708
Positive	34.6099	0.9657

TABLE 2: The experimental results.

Backbone	With LE	Accuracy	Precision	Recall	F1-score	Kappa
VGG-19	✓	82.8	76.82	94.47	84.73	65.5
		85.59	81.11	93.19	86.73	71.13
ResNet-50	✓	81.72	77.78	89.36	83.17	63.38
		83.23	78.97	91.06	84.59	66.39
DenseNet-121	✓	81.51	78.54	87.23	82.66	62.96
		83.01	82.77	83.83	83.3	66.01
Inception-v3	✓	82.15	78.36	89.36	83.5	64.24
		83.01	78.68	91.06	84.42	65.96

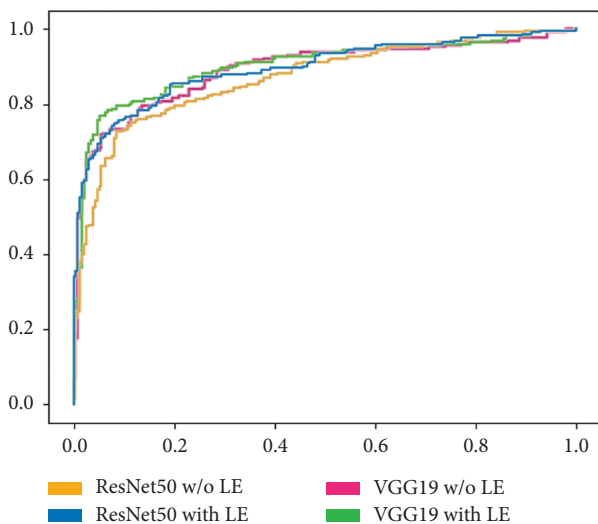


FIGURE 4: The ROC curve of ResNet and VGG.

words, compared to VGG with a simple straight-line connection, ResNet and other networks have a more complicated network structure, such as short connections. It may help the modeling process to better find the lesion area to a certain extent. However, more in-depth theoretical proofs and related experiments are required to prove this.

5. Conclusion

With the gradual improvement of deep neural network technology recently, there are more methods focusing their researches on exploring the hidden information inside data. Taking computer vision as an example, many studies have bundled the task of exploring the semantic information inside the image with the optimization of the neural network structure. Among them, medical imaging has become one of the important objects of image semantic mining due to its high semantic consistency. However, the interpretability of deep learning itself has always been a topic still under studying; that is, the characteristics learned by neural

networks cannot be explained intuitively. In other words, the correlation between the optimization of neural network structure and the effectiveness of semantic information mining remains to be verified. Therefore, this paper proposes an Adversarial Lesion Enhancement Neural Network for Medical Image Classification, which takes the extraction of this hidden semantics as a separate stage, separating it from the auxiliary diagnosis model. The purpose of this is to (1) clearly show the effectiveness of semantic information extraction, (2) use this method as a portable auxiliary diagnostic module with high adaptability, and (3) complete the lesion positioning with only coarse-grained labels. Finally, this paper proves the effectiveness of the first stage structure-based lesion adversarial inpainting module on the public dataset MURA. Meanwhile, on this basis, it is proved that the combined use of two-stage modules can improve the auxiliary diagnosis model. However, there is also a shortcoming in this method, that is, the relatively high time complexity of the sliding window, which is also one of the optimization directions of this research in the future.

Data Availability

The X-ray images used to support the findings of this study have been deposited in the MURA repository.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] I. Stuart, E. H. Y. Weinstein, and I. Sylvia, "Musculoskeletal research," 2014, <http://www.boneandjointburden.org/2014-report>.
- [2] D. Banga and W. Peter, "Abnormality detection in musculoskeletal radiographs with convolutional neural networks(ensembles) and performance optimization," *CoRR*, vol. 217, 2019.
- [3] S. Macias-Velasquez and A. Yolanda, "Working hours, burnout and musculoskeletal discomfort in middle and senior management of mexican industrial sector," *IEEE Access*, vol. 8, pp. 48607–48619, 2020.
- [4] K. Wang, L. Wang, Z. Deng, C. Jiang, W. Niu, and M. Zhang, "Influence of passive elements on prediction of intradiscal pressure and muscle activation in lumbar musculoskeletal models," *The Computer Methods and Programs in Biomedicine-Update*, vol. 177, pp. 39–46, 2019.
- [5] T. Tien-Loc Le, Lo-Yi Lin, C.-M. Lin, and F. Chao, "A k-means interval type-2 fuzzy neural network for medical diagnosis," *International Journal of Fuzzy Systems*, vol. 21, no. 7, pp. 2258–2269, 2019.
- [6] S. Raouafi, S. Achiche, M. Begon, A. Sarcher, and M. Raison, "Classification of upper limb disability levels of children with spastic unilateral cerebral palsy using k-means algorithm," *Medical & Biological Engineering & Computing*, vol. 56, no. 1, pp. 49–59, 2018.
- [7] M. Al-Maitah, "Analyzing genetic diseases using multimedia processing techniques associative decision tree-based learning and hopfield dynamic neural networks from medical images," *Neural Computing and Applications*, vol. 32, no. 3, pp. 791–803, 2020.

- [8] H. Faris, M. Habib, M. Faris, M. Alomari, and A. Alomari, "Medical speciality classification system based on binary particle swarms and ensemble of one vs. rest support vector machines," *Journal of Biomedical Informatics*, vol. 109, Article ID 103525, 2020.
- [9] C. Guo, J. Zhang, Y. Liu, Y. Xie, Z. Han, and J. Yu, "Recursion enhanced random forest with an improved linear model (RERF-ILM) for heart disease detection on the internet of medical things platform," *IEEE Access*, vol. 8, pp. 59247–59256, 2020.
- [10] M. Gadermayr, B. Mara Klinkhammer, B. Peter, and D. Merhof, "CNN cascades for segmenting sparse objects in gigapixel whole slide images," *Computerized Medical Imaging and Graphics*, vol. 71, pp. 40–48, 2019.
- [11] X. Cao, M. Liu, F. Zhai et al., "Comparison of different registration methods and landmarks for image-guided radiation therapy of pulmonary tumors," *BMC Medical Imaging*, vol. 19, no. 1, pp. 1–46, 2019.
- [12] P. Rajpurkar, J. Irvin, A. Bagul et al., "MURA: large dataset for abnormality detection in musculoskeletal radiographs," in *Proceedings of the 1st Conference on Medical Imaging with Deep Learning (MIDL 2018)*, Amsterdam, The Netherlands, 2018.
- [13] A. Amirkhani, M. R. Mosavi, K. Mohammadi, and E. I. Papageorgiou, "A novel hybrid method based on fuzzy cognitive maps and fuzzy clustering algorithms for grading celiac disease," *Neural Computing and Applications*, vol. 30, no. 5, pp. 1573–1588, 2018.
- [14] J. S. B. Sang Gil Lee, H. Kim, H. Jung, and S. Yoon, "Liver lesion detection from weakly-labeled multi-phase ct volumes with a grouped single shot multibox detector," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, 2018.
- [15] N. Dong, Li Wang, L. Xiang, S. Zhou, E. Adeli, and D. Shen, "Difficulty-aware attention network with confidence learning for medical image segmentation," in *Proceedings of the The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019*, pp. 1085–1092, AAAI Press, Honolulu, Hawaii, USA, 2019.
- [16] X. Wang, Y. Peng, Le Lu, Z. Lu, M. Ronald, and Summers, "Tienet: text-image embedding network for common thorax disease classification and reporting in chest x-rays," in *Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018*, pp. 9049–9058, IEEE Computer Society, Salt Lake City, UT, USA, 2018.
- [17] Qi Wang, J. Qiu, Y. Zhou, R. Tong, D. Gao, and Ju Gao, "Automatic severity classification of coronary artery disease via recurrent capsule network," in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2018*, pp. 1587–1594, IEEE Computer Society, Madrid, Spain, 2018.
- [18] P. Kumar 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 (CICT)*, Allahabad, India, 2019.
- [19] P. Rajpurkar, J. Irvin, K. Zhu et al., "Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning," *CoRR*, vol. 225, 2017.
- [20] A. F. M. Saif, C. Shahnaz, W.-P. Zhu, and M. Omair Ahmad, "Abnormality detection in musculoskeletal radiographs using capsule network," *IEEE Access*, vol. 7, pp. 81494–81503, 2019.
- [21] L. Oakden-Rayner, J. Dunnmon, G. Carneiro, and R. é Christopher, "Hidden stratification causes clinically meaningful failures in machine learning for medical imaging," in *Proceedings of the ACM CHIL '20: ACM Conference on Health, Inference, and Learning*, pp. 151–159, Toronto, Canada, 2020.
- [22] Y. Ren, X. Yu, R. Zhang, T. H. Li, S. Liu, and Li Ge, "Structureflow: image inpainting via structure-aware appearance flow," in *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019*, pp. 181–190, IEEE, Seoul, Korea, 2019.
- [23] Li Xu, Q. Yan, Y. Xia, and J. Jia, "Structure extraction from texture via relative total variation," *ACM Transactions on Graphics*, vol. 31, no. 6, p. 139, 2012.

Research Article

Nonlinear Load Harmonic Prediction Method Based on Power Distribution Internet of Things

Yongle Dong,¹ Fan Zhang,¹ Xuan Li,¹ Lifang Zhang,¹ Jia Yu,¹ Yongmei Mao,¹
and Guanglong Jiang ²

¹Inner Mongolia Electric Power Science & Research Institute, Hohhot, Inner Mongolia 010051, China

²Hexing Electrical Co., Ltd, Hangzhou, Zhejiang 310011, China

Correspondence should be addressed to Guanglong Jiang; guanglong.jiang@hxgroup.com

Received 8 March 2021; Revised 12 April 2021; Accepted 12 May 2021; Published 25 May 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Yongle Dong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A large number of nonlinear loads have an impact on the stable operation of the power system. To solve this problem, this article proposes a nonlinear load harmonic prediction method based on the architecture of Power Distribution Internet of Things. Firstly, this method integrates the characteristics of edge computing technology and Power Distribution Internet of Things technology and proposes a Power Distribution Internet of Things framework applied to nonlinear load harmonic prediction, which provides top-level design for subsequent harmonic prediction methods of Power Distribution Internet of Things; then, considering the electrical characteristics of the typical nonlinear load, the mathematical model of nonlinear load data is constructed based on the harmonic coupling admittance matrix model on the edge side. At the same time, a nonlinear load harmonic prediction model based on dynamic time warping and long-term and short-term memory network (DTW-LSTM) is established in the cloud computing center to realize high accuracy and high real-time prediction and analysis of nonlinear load harmonics. Finally, the simulation results based on the general data set show that the MAE evaluation index of the proposed method is less than 5% in the experimental group, which shows good generalization ability, and has some advantages over the current method in operation efficiency.

1. Introduction

With the rapid development of electrical technology and the electronic manufacturing industry, power electronic rectifying devices with nonlinear characteristics have been widely used in household appliances [1–3], such as energy-saving lamps, battery car chargers, and other electrical appliances, the harmonic distortion rate can reach more than 100%, and the harmonic distortion rate of the low-voltage side of the residential distribution network is nearly 20%, which makes the residential load become a new type of harmonic source [4].

As a small power nonlinear load, household appliances have the characteristics of large number and wide distribution and are an important harmonic component in residential distribution network [5–7]. The measured data show that the total demand distortion coefficient of harmonic current on some residential distribution feeders can

reach 12%, which makes the residential load become a very important harmonic source [8]. The harmonics generated by harmonic source will lead to the increase of harmonic loss in the residential distribution network, overload of transformer and neutral line, increase of failure probability of electronic equipment, and other problems, which will cause great difficulties in power quality evaluation and stable operation of distribution network [9–11]. Therefore, in order to ensure the sustainable and stable energy supply of distribution network side, it is necessary and urgent to study an efficient and reliable nonlinear load harmonic prediction method.

The traditional load harmonic prediction method adopts a linear modeling method, but it ignores the nonlinear and time-varying characteristics of the actual load [12, 13] and can not describe the nonlinear load harmonic variation law, which makes the prediction result deviate greatly.

Thanks to the continuous promotion and development of the Power Distribution Internet of Things, edge computing technology has made good progress in distribution network situation awareness and operation control [14–16]. In the Power Distribution Internet of Things, the distribution cloud master station can decentralize certain computing power to the edge side intelligent distribution terminal [17], realizing the preliminary processing of user state data at the edge side; in the distribution cloud master station, based on big data technology and artificial intelligence technology [18, 19], through iterative training and learning of multilayer network model, accurate state analysis and decision control of power system can be realized. Therefore, the combination of Power Distribution Internet of Things and artificial intelligence technology can effectively improve the performance of harmonic prediction and analysis for nonlinear load in the distribution system.

2. Related Work

Nonlinear load is an unstable source in power grid; when a large number of harmonic currents are injected into the power grid, it will cause voltage flicker, frequency fluctuation, three-phase voltage, and current imbalance, which seriously affect the transmission efficiency and the operation safety of equipment [20, 21]. Scientific and reasonable prediction and analysis of nonlinear load harmonics can effectively help power companies to formulate corresponding effective transmission strategies and ensure the stable and efficient operation of the distribution system [22]. At present, academic and industrial circles have given full attention to nonlinear load harmonic prediction.

Traditional harmonic load forecasting is mainly based on the circuit structure and electrical parameters of power load for mathematical modeling and analysis, such as the equivalent circuit analysis method [23, 24]. The equivalent circuit modeling and analysis method is mainly used to solve the analytical relationship between harmonic voltage and harmonic current by determining the simplified equivalent circuit of load and then deduce the equivalent harmonic model of load. However, it should be noted that the user load in the actual scene has the characteristics of diversity and complexity, so the power supply circuit structure is complex, which makes the equivalent circuit difficult to obtain and the circuit parameters difficult to estimate. These problems also lead to difficulties in modeling and low prediction accuracy in circuit analysis.

As a product of artificial intelligence technology, big data analysis technology plays a certain role in applying big data decision-making processing, such as power grid operation status monitoring and power transmission and distribution control [25]. Meanwhile, for the distribution network nonlinear load harmonic prediction, its essence is also for the power grid to collect big data to realize calculation and analysis, so some researchers have carried out some research.

In [26], considering the time series characteristics of power grid data, a new method of power grid harmonic prediction and analysis based on long-term and short-term

memory network model is proposed. Through multilevel network training and fitting learning of power grid state data, the efficient analysis of power quality is realized. In [27], the load current harmonics injected into the microgrid are predicted based on the Nonlinear AutoRegressive neural networks with eXogenous input (NARX), and the current harmonics caused by the nonlinear load are identified and isolated by the network training and learning of the measured grid data. In [28], a load harmonic analysis method based on a data-driven neural network is proposed. Through the analysis and modeling of load data in the time domain and frequency domain, it can effectively realize the research and analysis of nonlinear load long-term forecasting. In [29], in order to solve the problem of low load forecasting accuracy caused by the nonlinear characteristics of power load, an improved support vector machine model based on empirical mode decomposition and genetic algorithm is proposed to realize accurate load forecasting.

In view of the existing load forecasting research work, this article proposes a nonlinear load harmonic forecasting method based on the Power Distribution Internet of Things architecture. The main contents of this article are as follows:

- (1) The traditional centralized computing mode of distribution network with the main station as the core has the problem of low efficiency. This article constructs a “cloud, edge, end” three-tier architecture of Power Distribution Internet of Things combined with edge computing technology. Based on the demand of power consumption behavior analysis, the load harmonic prediction analysis framework in the Power Distribution Internet of Things is further proposed, and the data transmission, data processing, and task control mechanisms in the proposed prediction framework are sorted out in turn, which provides complete framework support for the subsequent nonlinear load harmonic prediction method based on deep learning network.
- (2) In order to improve the real-time performance and accuracy of nonlinear load forecasting analysis in the distribution system, the harmonic coupling admittance matrix model based on measured data is used to build a mathematical model of user side load at the edge of Power Distribution Internet of Things to solve the problem of difficult modeling caused by complex electrical structure and parameters. Based on dynamic time warping and long-term and short-term memory network (DTW-LSTM) deep learning network model in the cloud side, the nonlinear load harmonic prediction and analysis can be realized, which can reduce the number of load prediction models to be constructed, make the prediction model have strong generalization ability, and realize the high-precision and low delay prediction of nonlinear load harmonic in the distribution system.

The rest of this article is organized as follows. The third section introduces the collaborative architecture of distribution system, including the edge computing architecture of

Power Distribution Internet of Things and the nonlinear load harmonic prediction architecture. In the fourth section, the nonlinear load harmonic prediction method based on deep learning is introduced, including the mathematical modeling of nonlinear load harmonic electrical characteristics on the edge side and the load harmonic prediction method based on DTW-LSTM network model on the cloud side. The fifth section introduces the simulation analysis of the feasibility of the proposed method based on the EUNITE network dataset. The sixth section is the conclusion of this article.

3. System Architecture

3.1. Edge Computing Architecture of Power Distribution Internet of Things. Thanks to the concept and technology of edge computing, Power Distribution Internet of Things with “cloud, edge and terminal” three-tier architecture is constructed, as shown in Figure 1.

Among them, “cloud” refers to the cloud-based power distribution master station platform, that is, the distribution cloud master station, which has a variety of global decision-making services deployed on the cloud. “Edge” refers to the edge power distribution device, which is close to the end side equipment or data source, and provides edge intelligent services nearby. There are many microapplications deployed in it (based on the intelligent terminal, the application program that realizes specific business functions through independent development of software, imitating the concept of smartphone). “Terminal” is the main body of state aware and executive control in the Power Distribution Internet of Things architecture, which can monitor, collect, and perceive the basic data such as the operation environment, equipment status, and electrical quantity information of distribution equipment.

Edge computing technology is the core link of the three-tier architecture of power distribution networking. It is an open platform between data collection, computation, and application integration between intelligent terminal and distribution cloud master station. It is the carrier and key link of “intelligent terminal self-organization and terminal cloud self coordination.” With “edge cloud collaboration and edge intelligence” as the core feature, “edge cloud and cloud gateway” as the main landing form, and “software micro application” as the implementation mode, cloud computing is the extension and evolution of the sink node outside the data center, with the functions of collection, communication, calculation, analysis, and control. Through the deployment of microapplications in the edge power distribution device, we can flexibly upgrade and expand the terminal distribution business functions, make full use of the edge computing architecture advantages of local computing, develop high-value microapplications, and reflect the application value of the distribution Internet of things.

Referring to the functional architecture of industrial Internet platform, in order to better serve the power distribution industry, the edge layer is further designed into

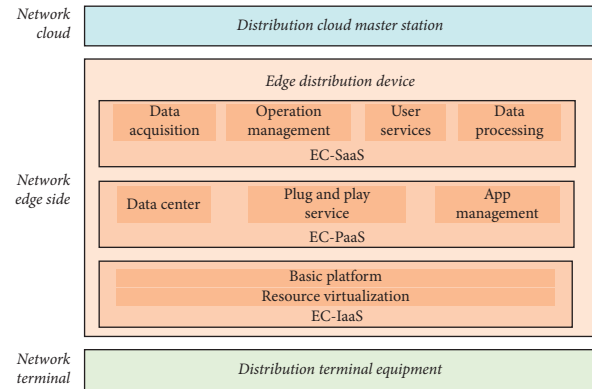


FIGURE 1: The architecture of Power Distribution Internet of Things.

three layers: Edge Computing Infrastructure as a Service (EC-IaaS), Edge Computing Software as a Service (EC-SaaS), and Edge Computing Platform as a Service (EC-PaaS), and the edge power distribution device is taken as the platform carrier.

EC-IaaS layer is the platform layer, which is the basic open platform of edge computing nodes, including hardware platform, operating system, container open platform, communication open platform, and AI engine. It provides unified computing, storage, communication, and system service capabilities for edge intelligent distribution business.

EC-PaaS layer is the software layer, which provides a backplane for all kinds of software operation, realizes data interaction and management, and supports application management monitoring, microapplication services, messages and events, data storage, and management software. At the same time, in order to meet the technical requirements of the “plug and play” of distribution network equipment, the plug and play service is used as the software layer to provide the basis for other applications.

EC-SaaS layer is the application layer, which is a microapplication service developed and deployed according to the demand of the distribution business. It is a specific way to implement the edge computing technology of distribution network. Through data collection or data cloud, end-to-end collaboration meets the operation and maintenance and power demand and provides data agent services for data interaction.

The edge power distribution device on the edge side mainly consists of two functional layers, namely, the management layer and the security layer. The management layer of the edge power distribution device is responsible for supporting remote and local software upgrade, user setting, password policy configuration, log audit configuration, management configuration of the edge computing node, and supporting system status monitoring and query. The security layer is responsible for controlling the access rights of system users, controlling the local or remote access of data, and verifying the legitimacy and integrity of the data source of the update package during software upgrade.

3.2. Nonlinear Load Harmonic Prediction Framework.

Based on the analysis of the network characteristics of the current distribution network, we can see that the amount of power distribution equipment is quantified, and the distribution monitoring data is also high-dimensional. This requires efficient edge devices to realize data preprocessing before uploading to the cloud, which can ensure real-time and efficient analysis of power consumption behavior in the cloud, so as to support the steady-state operation of the increasingly large Internet of Things. This coincides with the analysis concept of “cloud, edge, end” architecture of Power Distribution Internet of Things proposed in the previous content.

The specific network model architecture of nonlinear load harmonic prediction analysis framework is shown in Figure 2. The load forecasting model of Power Distribution Internet of Things is divided into three layers: distribution Infrastructure layer on the terminal side, intelligent distribution equipment layer on the edge side, and distribution cloud master station.

Among them, the network terminal side of the Power Distribution Internet of Things includes the power consumption equipment and monitoring equipment in the distribution network, which mainly involves the collection and transmission of sensing data and load data near the user side. The network edge side of Power Distribution Internet of Things is set as the edge network facilities of smart grid, such as smart substation, which carries out centralized preprocessing for the data uploaded by power terminal equipment. The cloud is the data center, which can carry out large-scale load harmonic prediction and analysis in its intelligent algorithm to support the steady-state operation of an intelligent distribution network.

In order to better support the harmonic prediction of nonlinear load in the distribution network, this article will elaborate and analyze three mechanisms of distribution network data transmission, data processing, and task control:

- (1) Data transmission mechanism: data transmission mainly focuses on data acquisition and screening of load forecasting. A large number of IoT sensing devices are deployed in the power distribution IoT system [30]. Therefore, these sensing devices can transmit the temperature, humidity, weather conditions, and other data to the edge power distribution device to achieve multitime and multistate load data recording. The essence of an edge power distribution device is an edge computing node, which not only undertakes the function of collecting sensing data but also can sense and classify the content of data. The edge power distribution device transmits the processed data and the data that need to be transferred to the cloud side to the distribution cloud master station and realizes the task of big data storage and data operation on the cloud side. At the same time, the distribution cloud master station will also transmit the historical data needed by the edge side calculation to the edge power distribution device.
- (2) Data processing mechanism: data processing mainly studies the data processing of load forecasting in the Power Distribution Internet of things under the cloud side architecture. The direct prediction of the load at any time will cause the number of neurons in the hidden layer to be too large, resulting in reduced computation efficiency. Therefore, a short-term load forecasting method based on deep learning is proposed. In the edge computing environment, the data processing tasks with low energy consumption and high delay requirements are deployed on the edge side of the Power Distribution Internet of Things (such as mathematical modeling of load electrical characteristics), and the data processing tasks with high energy consumption and low delay requirements are deployed on the cloud side of the main station (such as load forecasting and analysis behavior). Based on the cloud edge collaborative processing method, the computing tasks are distributed to the edge power distribution device, which can not only reduce the data transmission but also make the edge power distribution device closer to renewable energy, which can achieve more efficient and accurate load forecasting analysis.
- (3) Task control mechanism: task control mainly focuses on the scene of load forecasting task of edge power distribution devices and distribution cloud master station in the Power Distribution Internet of Things. By scheduling the load forecasting tasks, the delay and power consumption of the forecasting system can be reduced to improve the performance of the forecasting system.

4. Nonlinear Load Harmonic Prediction Method Based on Deep Learning

In order to achieve high accuracy harmonic prediction for user's nonlinear load, firstly, based on the admittance matrix model, the mathematical model of harmonic electrical characteristics of nonlinear load is established at the edge side. Then, based on the harmonic prediction architecture of Power Distribution Internet of Things, the deep learning algorithm is used to achieve efficient prediction and analysis of user's load harmonic at the cloud side.

4.1. Modeling of Harmonic Electrical Characteristics of Edge Side Nonlinear Load

4.1.1. Harmonic Characteristic Analysis of Nonlinear Load.

In order to study the harmonic generation of nonlinear loads of household appliances, the data of 21 loads and 28 different operation modes are measured, as shown in Table 1.

In order to classify the measured household loads, total harmonic current distortion (ITHD) and rate of harmonic decay (RHD) are used as the measurement indexes. The calculation formula of RHD is as follows:

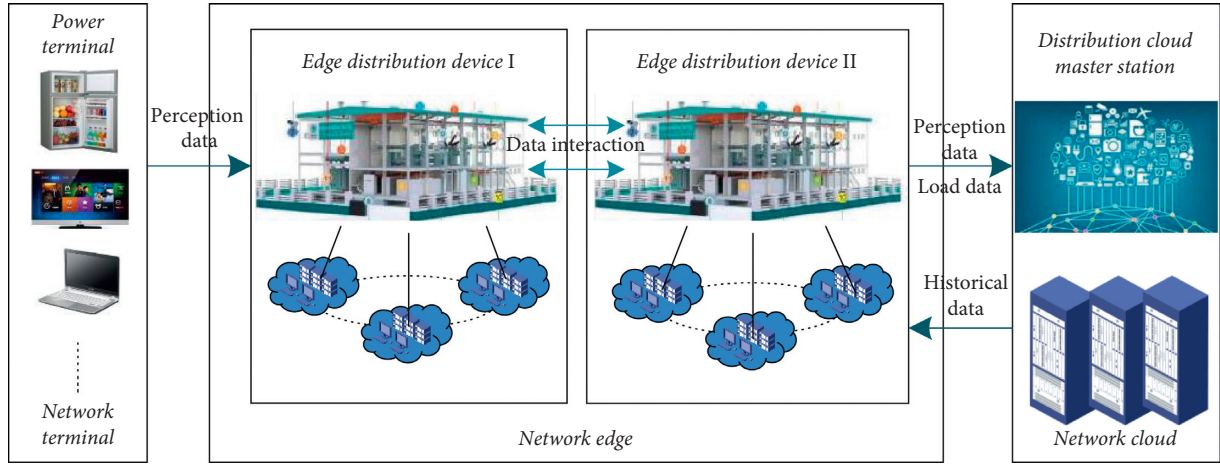


FIGURE 2: Nonlinear load harmonic prediction framework.

$$I_h = \frac{I_1}{h^{RHD}}, \quad (1)$$

where I_h is the h -th harmonic current amplitude; I_1 is the fundamental current amplitude; h^{RHD} is the harmonic attenuation rate of the h -th harmonic. Due to the variety and irregularity of current waveforms of household nonlinear load, this article obtains the RHD of each load based on the measured data of load.

The load of household appliances is further classified as follows: (1) power electronic household load with single-phase rectifier bridge (RHD is 0.32 ~ 0.8, current waveform is generally spike pulse shape); (2) the current waveform does not have typical characteristics (RHD is 0.8 ~ 1.2, the current waveform does not have typical characteristics); (3) the motor household load supplied by single-phase rectifier bridge (RHD is 1.2 ~ 2.73, current waveform is similar to triangle wave). For multistate nonlinear load, each state can be modeled according to the actual data.

4.1.2. Mathematical Model. Based on the above analysis, it can be seen that different types of nonlinear household loads have large differences, and it is difficult to model the nonlinear load data by using the circuit model analysis method because of the complex electrical structure and the difficulty in unifying the parameters.

In the past, the constant current source model was used to simplify the influence of the supply voltage on the load harmonic current so that the spectrum of the harmonic source cannot follow the change of the supply voltage. Therefore, the harmonic load sensitive to the supply voltage will produce large errors, especially the load of household appliances with rectifier and inverter. The higher the distortion rate of harmonic current is, the more serious the load is affected by the distortion of the supply voltage. Therefore, in order to accurately evaluate harmonics, a harmonic coupled admittance matrix (HCAM) model is proposed,

TABLE 1: Harmonic characteristics of nonlinear load for residential users.

Load	ITHD/%	RHD
Agricultural machinery washing	170	0.40
Washing machine	48	1.31
Drying of washing machine	81	0.61
Refrigerator insulation	23	1.49
Refrigerator refrigeration	21	1.49
Refrigerator defrosting	129	0.51
Air conditioning heating	88	0.91
Air conditioning insulation	21	1.48
Air conditioning refrigeration	26	1.91
Microwave oven rotation	11	2.14
Microwave oven heating	42	1.02
Bread machine	68	0.89
Router	171	0.39
Fluorescent lamp	81	0.66
LED light	71	0.64
CFL	92	0.71
TV set top box	139	0.53
Floor sweeping robot	131	0.51
Air cleaner	146	0.49
Air humidifier	139	0.55
Notebook computer	151	0.41
Desktop computer	43	0.87
Television	11	1.74
CRT TV	141	0.49
Plasma TV	4	2.61
Water purifier	16	1.71
Battery car charger	121	0.66
Treadmill	162	0.51

which can consider the influence of supply voltage fluctuation and harmonic distortion.

Compared with the traditional constant current source model or Norton model, the HCAM model can reflect the harmonic generation of load more accurately. The specific form of the HCAM model is shown in

$$\begin{bmatrix} \dot{I}_1 \\ \dot{I}_3 \\ \vdots \\ \dot{I}_i \\ \vdots \\ \dot{I}_H \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{13} & \cdots & Y_{1j} & \cdots & Y_{1H} \\ Y_{31} & Y_{33} & \cdots & Y_{3j} & \cdots & Y_{3H} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{i1} & Y_{i3} & \cdots & Y_{ij} & \cdots & Y_{iH} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{H1} & Y_{H3} & \cdots & Y_{Hj} & \cdots & Y_{HH} \end{bmatrix} \begin{bmatrix} \dot{V}_1 \\ \dot{V}_3 \\ \vdots \\ \dot{V}_j \\ \vdots \\ \dot{V}_H \end{bmatrix} = YV, \quad (2)$$

where \dot{I}_i and \dot{V}_j are the phasor values of the i -th harmonic current and the j -th harmonic voltage, respectively; Y_{ij} represents the contribution of the j -th harmonic voltage to the i -th harmonic current; $i = 1, 2, \dots, H$, $j = 1, 2, \dots, H$ and the highest harmonic is H .

HCAM model can accurately represent the harmonic coupling characteristics of nonlinear load, but the structure and parameters of the load equivalent circuit need to be known for modeling. However, most uncertain load equivalent circuit structures are complex and parameters are unknown, so the model elements of the harmonic coupling matrix can be estimated by using measured power consumption data.

In the process of establishing the model, it is found that because the Y matrix is the full rank matrix, and there are many model elements and multiple collinearity, which leads to matrix ill-conditioned phenomenon in the process of solving element parameters. In order to overcome this problem, based on the amplitude characteristics of matrix elements, it is found that the coupling between harmonic voltage and current with the same harmonic number and adjacent ones is the strongest. Therefore, it is proposed to consider only the first column elements of Y matrix, the main diagonal elements, and the adjacent elements of the same row of main diagonal elements, as shown in the following formula:

$$\begin{bmatrix} \dot{I}_1 \\ \dot{I}_3 \\ \vdots \\ \dot{I}_h \\ \vdots \\ \dot{I}_H \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{13} & 0 & \cdots & \cdots & 0 \\ Y_{31} & Y_{33} & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ Y_{i1} & 0 & Y_{h(h-2)} & Y_{hh} & Y_{h(h+2)} & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{H1} & 0 & \cdots & 0 & Y_{H(H-2)} & Y_{HH} \end{bmatrix} \begin{bmatrix} \dot{V}_1 \\ \dot{V}_3 \\ \vdots \\ \dot{V}_h \\ \vdots \\ \dot{V}_H \end{bmatrix}. \quad (3)$$

Among them, the first column element is the interaction between fundamental voltage and each harmonic current, and the main diagonal element is the interaction between each harmonic voltage and the same harmonic current or two adjacent harmonic currents. Equation (3) can be further divided into the effect of fundamental voltage and the effect

of each harmonic voltage. As shown in equation (4), it is the proposed nonlinear load harmonic coupling dominant component model.

$$\begin{bmatrix} \dot{I}_1 \\ \dot{I}_3 \\ \vdots \\ \dot{I}_h \\ \vdots \\ \dot{I}_H \end{bmatrix} = \begin{bmatrix} \dot{I}_{S1} \\ \dot{I}_{S3} \\ \vdots \\ \dot{I}_{Sh} \\ \vdots \\ \dot{I}_{SH} \end{bmatrix} + \begin{bmatrix} Y_{13} & 0 & \cdots & \cdots & 0 \\ Y_{33} & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & Y_{h(h-2)} & Y_{hh} & Y_{hh} & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & Y_{H(H-2)} & Y_{HH} \end{bmatrix} \begin{bmatrix} \dot{V}_3 \\ \vdots \\ \dot{V}_h \\ \vdots \\ \dot{V}_H \end{bmatrix}. \quad (4)$$

The current source represents the interaction between the fundamental voltage and the h -th harmonic current. Using the measured voltage and current of load, the current source and Y matrix elements in the model are obtained. For the fundamental parameters, by taking the load voltage and current at multiple times, we can get the following results:

$$\begin{bmatrix} \dot{I}_1(t_1) \\ \dot{I}_1(t_2) \\ \vdots \\ \dot{I}_1(t_m) \end{bmatrix} = \begin{bmatrix} 1 & \dot{V}_3(t_1) \\ 1 & \dot{V}_3(t_2) \\ \vdots & \vdots \\ 1 & \dot{V}_3(t_m) \end{bmatrix} \begin{bmatrix} \dot{I}_{S1} \\ Y_{13} \end{bmatrix} \Rightarrow \dot{I}_1 = B_1 X_1, \quad (5)$$

where t_1, t_2, \dots, t_m is the different voltage and current data of m groups.

The third harmonic parameters in the model can be calculated according to

$$\begin{bmatrix} \dot{I}_3(t_1) \\ \dot{I}_3(t_2) \\ \vdots \\ \dot{I}_3(t_m) \end{bmatrix} = \begin{bmatrix} 1 & \dot{V}_3(t_1) & \dot{V}_5(t_1) \\ 1 & \dot{V}_3(t_2) & \dot{V}_5(t_2) \\ \vdots & \vdots & \vdots \\ 1 & \dot{V}_3(t_m) & \dot{V}_5(t_m) \end{bmatrix} \begin{bmatrix} \dot{I}_{S3} \\ Y_{33} \\ Y_{35} \end{bmatrix} \Rightarrow \dot{I}_3 = B_3 X_3. \quad (6)$$

The calculation of the h' ($5 \leq h' \leq (H-2)$)-th harmonic parameter can be obtained by

$$\begin{bmatrix} \dot{I}_{h'}(t_1) \\ \dot{I}_{h'}(t_2) \\ \vdots \\ \dot{I}_{h'}(t_m) \end{bmatrix} = \begin{bmatrix} 1 & \dot{V}_{(h'-2)}(t_1) & \dot{V}_{h'}(t_m) & \dot{V}_{(h'+2)}(t_1) \\ 1 & \dot{V}_{(h'-2)}(t_2) & \dot{V}_{h'}(t_m) & \dot{V}_{(h'+2)}(t_2) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \dot{V}_{(h'-2)}(t_m) & \dot{V}_{h'}(t_m) & \dot{V}_{(h'+2)}(t_m) \end{bmatrix} \begin{bmatrix} \dot{I}_{Sh'} \\ Y_{h'(h'-2)} \\ Y_{h'h'} \\ Y_{h'(h'+2)} \end{bmatrix} \Rightarrow \dot{I}_{h'} = B_{h'} X_{h'}. \quad (7)$$

The highest order harmonic parameters can be obtained by

$$\begin{aligned} \begin{bmatrix} \dot{I}_H(t_1) \\ \dot{I}_H(t_2) \\ \vdots \\ \dot{I}_H(t_m) \end{bmatrix} &= \begin{bmatrix} 1 & \dot{V}_{(H-2)}(t_1) & \dot{V}_H(t_1) \\ 1 & \dot{V}_{(H-2)}(t_2) & \dot{V}_H(t_2) \\ \vdots & \vdots & \vdots \\ 1 & \dot{V}_{(H-2)}(t_m) & \dot{V}_H(t_m) \end{bmatrix} \\ &\cdot \begin{bmatrix} \dot{I}_{SH} \\ Y_{H(H-2)} \\ Y_H \end{bmatrix} \Rightarrow \dot{I}_H = B_H X_H. \end{aligned} \quad (8)$$

In combination with equations (5) to (8), the h -th ($1 \leq h \leq H$) harmonic parameters are obtained by using the least square method:

$$X_h = (B_h B_h^T)^{-1} B_h^T \dot{I}_h. \quad (9)$$

The load harmonic current calculated by harmonic coupling dominant component model can vary with the fluctuation of supply voltage, and the influence of supply voltage fluctuation and distortion on load harmonic current is considered; Moreover, the HCAM model intuitively reflects the coupling effect between harmonic voltage and harmonic current and more accurately reflects the harmonic generation characteristics of load, which can be used for the aggregation analysis of harmonic current when multiple harmonic source loads work together.

At the same time, the constant power model of linear and nonlinear load is established under the fundamental frequency by using the multiplier principle.

In this article, the exponential model is used to calculate the fundamental power, as shown in

$$\begin{aligned} P &= P_0 \left(\frac{V}{V_0} \right)^{n_p}, \\ \text{s.t. } n_p &\approx \frac{2 \times Z_p + 1 \times I_p + 0 \times P_p}{Z_p + I_p + P_p}, \end{aligned} \quad (10)$$

where P is the load active power when the voltage is V ; P_0 is the rated active power of the load; V_0 is the rated voltage amplitude of the system; Z_p , I_p , and P_p are the constant parameter to be solved.

Considering the influence of harmonic voltage variation and the coupling relationship between voltage and current on power, based on the above harmonic model and equation (10), the real-time harmonic power of residential load can be obtained as follows:

$$\begin{cases} P_h(t) = \sum_{h=1}^H U_h(t) I_h(t) \cos(\varphi_h(t)), Q_h(t) \\ = \sum_{h=1}^H U_h(t) I_h(t) \sin(\varphi_h(t)), \end{cases} \quad (11)$$

where $U_h(t)$, $I_h(t)$, and $\varphi_h(t)$, respectively, represent the h -th harmonic voltage, harmonic current, and phase angle difference between harmonic voltage and current of load at time t and H represents the highest harmonic considered.

4.2. Nonlinear Load Harmonic Prediction Based on Deep Learning in Cloud Side. Due to certain time series characteristics of power consumption data in the distribution network, in order to better mine the characteristics of nonlinear load data in the distribution system, this article uses the improved long-term and short-term memory neural network based on dynamic time warping to train and learn the load electrical characteristics data in the cloud side of Power Distribution Internet of Things and then realizes accurate and efficient nonlinear load harmonic prediction.

4.2.1. Dynamic Time Warping. Dynamic time warping (DTW) is a method to measure the similarity of two time series with different lengths [31]. DTW algorithm can ensure that when a peak value of the curve moves forward for a small period, it can still effectively find its similarity through the extension of time dimension, thus reducing the number of load forecasting models to be constructed.

Assume that the load electrical characteristic series of two users are, respectively, X and Y , where $X = \{x_1, x_2, \dots, x_N\}$, $Y = \{y_1, y_2, \dots, y_N\}$. Define the regular path matrix M [$24 * 24$], matrix element (m, n) represents the distance between x_m and y_n , and the value of matrix element (m, n) is $d(x_m, y_n) = (x_m - y_n)^2$. The total cost of dynamic regularization between load series X and Y is as follows:

$$c_p(X, Y) = \sum_{k=1}^K d(x_{m_k}, y_{n_k}). \quad (12)$$

Definition 1. Dynamic regularization path sequence $P = (p_1, p_2, \dots, p_k)$, where $p_k = (m_k, n_k)$ and $\max(m, n) \leq K < m + n - 1$.

Definition 2. Dynamic time warping distance of X and Y .

$$\text{DTW}(X, Y) = c_p(X, Y), \quad (13)$$

where $P^* = \arg \min c_p(X, Y)$ takes the weekly average DTW distance of two users as the clustering distance in the lower clustering algorithm, then

$$\text{DTW}(X, Y) = \frac{1}{7} \sum_{t=0}^6 \text{DTW}(X_t, Y_t), \quad (14)$$

where X_t and Y_t , respectively, represent the load curve of the two users on day t .

Given a series of user daily load curves X and the number of clusters D , users are clustered into class D according to the DTW distance of load curve, and class d , belonging function C_d , and cluster center λ_d are obtained to minimize sum S_c within the cluster.

$$S_c = \sum_d \sum_{x \in C_d} \text{DTW}(X, \lambda_d). \quad (15)$$

Then, the same type of user data is pooled to establish a unified load forecasting model, which is helpful to enhance the generalization ability of the model and improve the

accuracy of the forecasting model. The specific steps of pooling method are as follows: Firstly, the ID tag is added to the user in the form of virtual variable; then the user data is divided into training set R and test set T ; finally, all the training data are merged to build the training pool, and the test pool is built through the same process.

4.2.2. Long-Term and Short-Term Memory Neural Network.

The cloud side uses long-term and short-term memory (LSTM) model to predict the user's load for the new user data set after preprocessing. The internal structure of the LSTM neural network is shown in Figure 3. In order to establish time connection, LSTM defines and maintains an internal memory unit state cell state C_t in the whole cycle, and then updates, maintains, or deletes the information in the cell state through three gate structures: forgetting gate f_t , input gate i_t , and output gate o_t . The forward calculation process is as follows:

$$\begin{aligned}
 f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\
 i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\
 \tilde{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \\
 C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \\
 o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\
 h_t &= o_t \cdot \tanh(C_t),
 \end{aligned} \tag{16}$$

where C_t and C_{t-1} represent the cell state of the current time and the previous time, respectively, and C_t represents the candidate state of the input; f_t , i_t , and o_t , respectively, represent forgetting gate, input gate, and output gate; W_f , W_i , W_c , W_o , b_f , b_i , b_c , and b_o represent Sigmoid and hyperbolic tangent activation functions, respectively. Firstly, the coefficients of the forgetting gate, the input gate h_{t-1} , and the output gate x_t are calculated by the above formula; and the candidate states of the current neuron C_t are obtained by the formula of the last hidden layer output h_{t-1} and the current output x_t ; then, the proportion of the last hidden layer output C_{t-1} and the current hidden layer output C_t in the current cell state is determined by the forgetting gate and the input gate and the output value of the current hidden layer h_t is calculated by the formula.

The harmonic prediction model based on LSTM is trained by time backpropagation algorithm. The error term between the output value of each LSTM neuron and the real value is calculated. According to the corresponding error term, the gradient of each weight is calculated, and the weight is updated by the gradient optimization algorithm.

Due to the periodicity of power load data, the historical load data of the day before the load forecasting day, the week before the load forecasting day, and the month before the load forecasting day are weighted by a fully connected network and then used as the training dataset T of LSTM forecasting network.

In the test part, the test load curve is sent to the trained LSTM network. Suppose that the load curve data set after data cleaning is ΨI , and the test households are listed in the

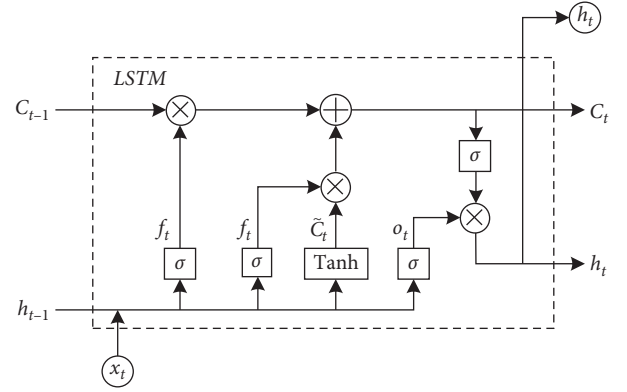


FIGURE 3: Internal structure of LSTM neural network.

set $D = \{d_1, d_2, \dots, d_d\}$. Then, we need to determine the network configuration parameters of LSTM, using L and H to represent the network depth (number of layers) and the number of hidden units, respectively. Through these parameters, the LSTM prediction network is initialized, and the network configuration parameters are established, i.e., network depth L , number of hidden layers H , and batch size C . Then, after the network is started, the program runs the training iteration period until the network is well trained.

In each training cycle, the training data are randomly selected from the training data pool and then input to the feedforward neural network for network training. Each training batch consists of two matrices with a fixed size, namely, input matrix with $C \times I$ size and output matrix with $C \times O$ size. The time cost and iteration times of the training process depend on the size of the feedback data sequence J , the optimization method selected, the size of the network (M, I) , and the size of the training batch C . In order to achieve a good balance between training efficiency and efficacy, the training batch size C is variable in the training process.

The specific flowchart of the distribution network nonlinear load harmonic prediction method based on the DTW-LSTM algorithm is shown in Figure 4.

5. Simulations and Results

In order to verify the feasibility and practicability of the proposed method, the simulation hardware environment is Lenovo Xiaoxin pro14, Intel Core i5 1135g7 four-core/eight-thread processor, 16 GB memory, NVIDIA Geforce mx450 independent graphics card; the software environment is Chinese Windows 10, English Microsoft Visual Studio 2012.

In this article, Caffe deep learning framework is used to test the open dataset of EUNITE network. EUNITE network dataset is the real power load data with sampling interval of 15 minutes, that is, 96 sampling points per day, a total of 28800 load points. The sampling time was from August 1, 2016, to August 7, 2016. The EUNITE network dataset is divided into training sample set and test sample set in the ratio of 3:1.

Before the algorithm test, this article preprocesses the load data to ensure the quality of the data, uses the mean

value to make up the missing value, and judges the abnormal value. At the same time, in the process of training, this article uses the Adam method to update the parameters. In order to prevent the model from overfitting, the dropout value is set to 0.5. Set the value of the maximum time step N_{\max} to 30.

5.1. Evaluation Index. Because the improved LSTM deep learning network harmonic prediction model is essentially a regression model, root mean square error (RMSE) and mean absolute error (MAE) are selected as the performance evaluation indexes based on the characteristics of the model. Among them, RMSE and MAE indicators represent the fitting deviation of the prediction model for the real parameter data, and the smaller the index value is, the more accurate the fitting result is, which can provide error visualization based on percentage. The calculation formula of RMSE and MAE is as follows:

$$\text{RMSE} = \frac{1}{n} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (17)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\%,$$

where n is the number of forecast points, y_i is the real load value of the i -th forecast point, and \hat{y}_i is the forecast value of the i -th forecast point.

5.2. Analysis of Experimental Results

5.2.1. DTW-LSTM Network Model Parameter Analysis. As an important parameter of the LSTM model, the value of the learning rate has an important influence on the effect of nonlinear load harmonic prediction. Therefore, it is necessary to determine the learning rate parameters of deep learning network model before load forecasting research and analysis. In this article, the training sample set data are used to optimize the learning rate parameters of the DTW-LSTM model, and the test results are shown in Figures 5 and 6.

From Figures 5 and 6, we can see that when the learning rate is 0.00015, the network jumps back and forth in the local optimum, the loss function and the accuracy of the model have obvious oscillation phenomenon, and the network cannot reach convergence after the set 200 iterations. On the contrary, when the learning rate is set to 0.0015, the network convergence speed is slow, the loss value and accuracy are not as good as when the learning rate is 0.015, the loss function value convergence is about 0.1 higher than when the learning rate is 0.015, and the model accuracy is reduced by 0.085. Obviously, the learning rate of 0.015 is the best, and the loss function converges to 0.168 when iterating to 23 times; the accuracy of the model converges to 0.964 after 23 iterations, and it is stable after 80 iterations, which can ensure that the network convergence is more stable and the convergence result is better.

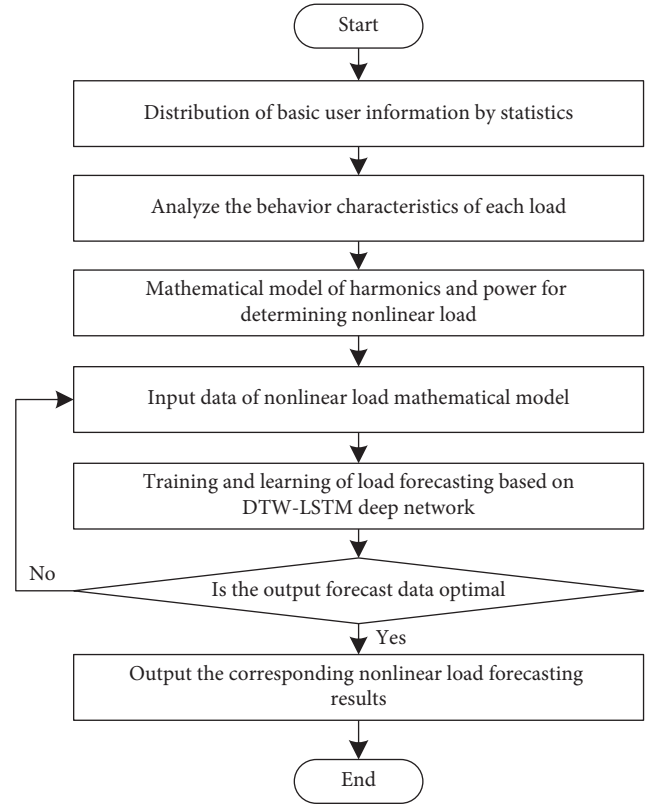


FIGURE 4: Flowchart of nonlinear load forecasting method based on deep learning network.

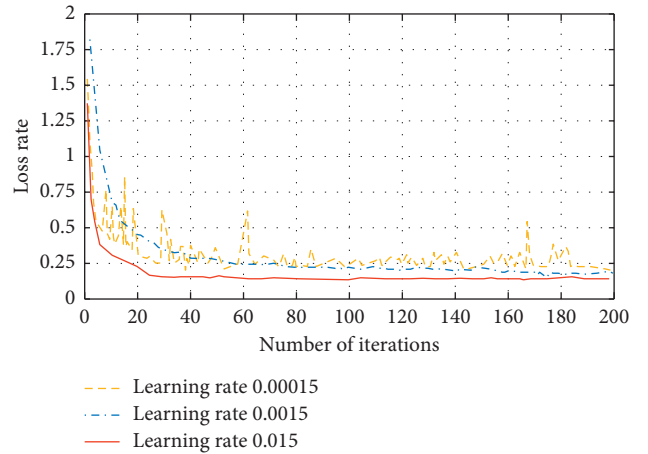


FIGURE 5: Loss rate of different learning rates.

5.2.2. Analysis of Harmonic Prediction Results. In order to verify the feasibility and superiority of the prediction method proposed in this article, [26], [27], and [29] are used as comparative methods to train and predict the EUNITE network dataset and the prediction accuracy of each method is analyzed based on the model evaluation indexes RMSE and MAE. Figure 7 shows the simulation results of different methods on the EUNITE network dataset.

It can be seen from Figure 7 that the DTW-LSTM prediction method proposed in this article has obvious

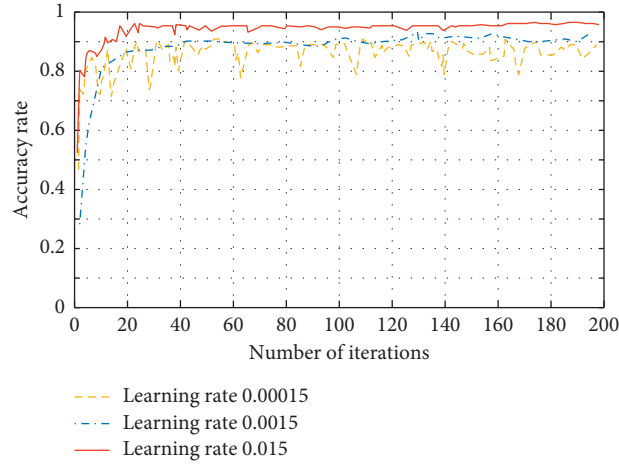


FIGURE 6: Accuracy of different learning rates.

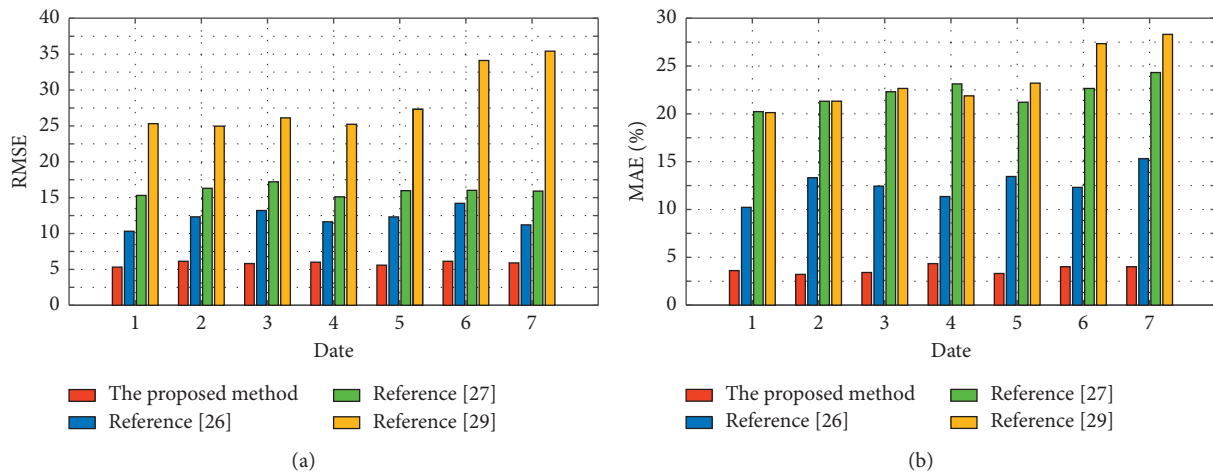


FIGURE 7: The prediction results under different methods: (a) RMSE and (b) MAE.

advantages over the other three methods in prediction performance. The seven-day average RMSE index obtained by this method is 5.84, which is 6.91, 10.06, and 22.53 higher than that of [26], [27], and [29], respectively. At the same time, the average value of MAE on the seventh day was 3.83%, and the maximum value of MAE on the sixth day was 4.87%, which was still within 5%. It is confirmed that the performance of the DTW-LSTM method presented in this article is stable in 7 consecutive days.

It should be noted that in the dataset used in this article, 6 and 7 days are weekend rest days, and the RMSE and Mae values of the proposed method in these two days are significantly different from those in the working days from August 1 to 5; However, using the prediction method of [29] to predict the nonlinear harmonics of rest days and working days, the evaluation index values differ greatly, which shows that the prediction method used in this article has good generalization ability.

5.2.3. Operation Efficiency Analysis. In order to analyze the efficiency of nonlinear harmonic load forecasting, the training time and test time of nonlinear harmonic load forecasting modeling are used as evaluation indexes to count the training time and test time of nonlinear harmonic load. The results are shown in Table 2.

From Table 2, we can see that the training time and running time of the proposed method for EUNITE Network load dataset are 114.22 s and 51.43 s, respectively, which is obviously faster than that of [26] and [27]. However, it is noted that the processing time of the training process of the sample dataset in [29] is 0.33 s shorter than that of the proposed DTW-LSTM method, and the processing time of the nonlinear load harmonic prediction experiment of the EUNITE network data set is 1.89 s longer than that of the proposed method. At the same time, combined with the above comparison of the prediction accuracy of the two methods based on the model evaluation indexes RMSE and Mae, it can be

TABLE 2: The running time under different methods.

Classification method	Experimental running time/s	
	Training dataset	Test dataset
The proposed method	114.22	51.43
Reference [26]	117.34	61.52
Reference [27]	120.32	73.53
Reference [29]	113.89	53.32

seen that the proposed method has obvious advantages in terms of stability and generalization, although the processing time is close to that of [29]. Therefore, our method is still superior to the existing methods.

To sum up, compared with other methods, the proposed nonlinear load harmonic prediction method based on Power Distribution Internet of Things architecture has better performance in load harmonic prediction of distribution system and has better load harmonic prediction results than the existing methods under the premise of ensuring certain processing efficiency.

6. Conclusion

In order to meet the demand of high accuracy and real-time for nonlinear load harmonic prediction and analysis in the distribution system, this article proposes a nonlinear load harmonic prediction method based on the combination of distribution Internet of things and deep learning network. With the support of the architecture of Power Distribution Internet of Things based on nonlinear load harmonic prediction, this method uses the harmonic coupling admittance matrix model based on measured data at the edge of Power Distribution Internet of Things to realize the mathematical modeling of user side load data and solves the modeling problems caused by complex electrical structure and parameters; At the same time, the DTW-LSTM algorithm is used to build the load harmonic prediction and analysis model on the cloud side of the network, which can effectively reduce the number of load prediction models, ensure the strong generalization ability of the prediction model, and realize the efficient analysis of nonlinear load harmonic in the distribution system. Finally, based on the general dataset, the high-efficiency performance of the proposed method for nonlinear load harmonic prediction is verified, which proves that the proposed method has obvious advantages in processing accuracy and speed.

There are still a few data errors or data missing in the existing historical load data, so future research direction should focus on the nonlinear load harmonic prediction analysis based on the blockchain technology in the Power Distribution Internet of things under the scenario of data uncertainty.

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

This work was supported by the Inner Mongolia Power (Group) Co., Ltd. Science and Technology Project funding (issue no. NDKX[2020] No. 29): "Research on the influence of nonlinear load on the accuracy of electric energy measurement (Project no. [2020-21])."

References

- [1] G. Chang, C. Hatziaodoniou, W. Xu et al., "Modeling devices with nonlinear Voltage-current Characteristics for harmonic studies," *IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1802–1811, 2004.
- [2] K. Agyeman, S. Han, and S. Han, "Real-time recognition non-intrusive electrical appliance monitoring algorithm for a residential building energy management system," *Energies*, vol. 8, no. 9, pp. 9029–9048, 2015.
- [3] A. Javadi, A. Hamadi, A. Ndtoungou, and K. Al-Haddad, "Power quality enhancement of smart households using a multilevel-THSeAF with a PR controller," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 465–474, 2017.
- [4] Y. Yang, S. Li, W. Li, and M. Qu, "Power load probability density forecasting using Gaussian process quantile regression," *Applied Energy*, vol. 213, no. 1, pp. 499–509, 2018.
- [5] T. Ding, H. Liang, and W. Xu, "An analytical method for probabilistic modeling of the steady-state behavior of secondary residential system," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2575–2584, 2017.
- [6] J. Jiang, Z. Wen, M. Zhao et al., "Series arc detection and complex load recognition based on principal component analysis and support vector machine," *IEEE Access*, vol. 7, no. 1, pp. 47221–47229, 2019.
- [7] J. Yong and B. Xiao, "Harmonic modeling and investigation on harmonic attenuation effect for typical single-phase nonlinear loads," *Proceedings of the CSEE*, vol. 34, no. 19, pp. 3210–3219, 2014.
- [8] B. Shan, X. Sun, J. Li et al., "Analysis on the China's electricity demand growth under the new economic norm," *Electric Power*, vol. 50, no. 1, pp. 19–24, 2017.
- [9] M. Armendariz, D. Babazadeh, D. Brodén, and L. Nordström, "Strategies to improve the voltage quality in active low-voltage distribution networks using DSO's assets," *IET Generation, Transmission & Distribution*, vol. 11, no. 1, pp. 73–81, 2017.
- [10] G. Ye, M. Nijhuis, V. Cuk et al., "Stochastic residential harmonic source modeling for grid impact studies," *Energies*, vol. 10, no. 3, pp. 1–21, 2017.
- [11] R. Mitra, A. K. Goswami, and P. K. Tiwari, "Voltage sag assessment using type-2 fuzzy system considering uncertainties in distribution system," *IET Generation, Transmission & Distribution*, vol. 11, no. 6, pp. 1409–1419, 2017.
- [12] L. Guo, Y. Jin, and K. Luo, "Improved double vector model predictive current control method for low loss grid connected inverter," *Power Automation Equipment*, vol. 039, no. 10, pp. 136–142, 2019.
- [13] M. Cui, J. Wang, Y. Wang, R. Diao, and D. Shi, "Robust time-varying synthesis load modeling in distribution networks considering voltage disturbances," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4438–4450, 2019.

- [14] J. Lim, I. Doh, and K. Chae, "Secure and structured IoT smart grid system management," *International Journal of Web and Grid Services*, vol. 13, no. 2, pp. 170–185, 2017.
- [15] H. Mortaji, S. H. Ow, M. Moghavvemi, and H. A. F. Almurib, "Load shedding and smart-direct load control using Internet of things in smart grid demand response management," *IEEE Transactions on Industry Applications*, vol. 53, no. 6, pp. 5155–5163, 2017.
- [16] Q. Wang, S. Shao, S. Guo, X. Qiu, and Z. Wang, "Task allocation mechanism of power Internet of things based on cooperative edge computing," *IEEE Access*, vol. 8, no. 1, pp. 158488–158501, 2020.
- [17] T. Yang, F. Zhai, Y. Zhao et al., "Interpretation and research prospect of ubiquitous power Internet of things [J]," *Power System Automation*, vol. 43, no. 13, pp. 9–20 + 53, 2019.
- [18] X. Li, L. Guo, D. Huang et al., "Review on Key Technologies of DC distribution network operation control," *High Voltage Technology*, vol. 45, no. 10, pp. 3039–3049, 2019.
- [19] J. Zhao, X. Xia, C. Xu et al., "Review on the application of new generation artificial intelligence technology in power system dispatching," *Power System Automation*, vol. 44, no. 24, pp. 1–10, 2020.
- [20] Y. Chen, Z. Huang, Z. Duan et al., "A four-winding inductive filtering transformer to enhance power quality in a high-voltage distribution network supplying nonlinear loads," *Energies*, vol. 21, no. 10, pp. 1–13, 2021.
- [21] S. K. Shukla, E. Koley, S. Ghosh et al., "Enhancing the reliability of six-phase transmission line protection using power quality informatics with real-time validation," *International Transactions on Electrical Energy Systems*, vol. 29, no. 9, pp. 1–21, 2019.
- [22] J. Olamaei, F. Narjabadifam, and A. Nazeri, "A novel control scheme for simultaneous elimination of imbalance, disturbance and voltage harmonic in power systems," *Journal of Central South University*, vol. 21, no. 4, pp. 1368–1375, 2014.
- [23] B. Tang, S. Ma, S. Lin et al., "Probability distribution of residential load collective harmonics based on cloud theory," *Power System Automation*, vol. 40, no. 3, pp. 39–45 + 151, 2016.
- [24] Z. Yin, Y. Chen, Z. Liu et al., "Harmonic load model control strategy based on measured data," *Power Grid Technology*, vol. 40, no. 10, pp. 3192–3198, 2016.
- [25] D. Zhang, X. Miao, L. Liu et al., "Research on big data technology development of smart grid," *Chinese Journal of Electrical Engineering*, vol. 35, no. 1, pp. 2–12, 2015.
- [26] Q. Liu, W. Yin, W. Hu et al., "Prediction of power harmonic monitoring data based on LSTM algorithm," *Power Capacitor and Reactive Power Compensation*, vol. 40, no. 05, pp. 139–145, 2019.
- [27] A. Y. Hatata and M. Eladawy, "Prediction of the true harmonic current contribution of nonlinear loads using NARX neural network," *Alexandria Engineering Journal*, vol. 57, no. 3, pp. 1509–1518, 2018.
- [28] C.-I. Chen and Y.-C. Chen, "A neural-network-based data-driven nonlinear model on time- and frequency-domain voltage-current characterization for power-quality study," *IEEE Transactions on Power Delivery*, vol. 30, no. 3, pp. 1577–1584, 2015.
- [29] M. C. Shui, L. Y. Wei, J. W. Liu et al., "A hybrid one-step-ahead time series model based on GA-SVR and emd for forecasting electricity loads," *Journal of Applied Science and Engineering*, vol. 20, no. 4, pp. 4467–4476, 2017.
- [30] A. Albayati, N. F. Abdullah, A. Abu-Samah et al., "A serverless advanced metering infrastructure based on fog-edge computing for a smart grid: a comparison study for energy sector in Iraq," *Energies*, vol. 13, no. 20, pp. 1–22, 2020.
- [31] X. Wang, F. Yu, W. Pedrycz, and L. Yu, "Clustering of interval-valued time series of unequal length based on improved dynamic time warping," *Expert Systems with Applications*, vol. 125, no. 1, pp. 293–304, 2019.

Research Article

An Efficient Early Frame Breaking Strategy for RFID Tag Identification in Large-Scale Industrial Internet of Things

Zhiyong He ¹ and Hanguang Luo ²

¹College of Electrical Engineering, Nanjing Vocational University of Industry Technology, Nanjing 210046, China

²Research Center for Industrial Internet, Zhejiang Lab, Hangzhou 311121, China

Correspondence should be addressed to Hanguang Luo; luohg@zhejianglab.com

Received 16 April 2021; Accepted 12 May 2021; Published 20 May 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Zhiyong He and Hanguang Luo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the increase in the number of tags, an efficient approach of tag identification is becoming an urgent need in Industrial Internet of Things (IIoT). However, the identification performance of existing Aloha-based anticollision schemes is limited when the initial frame size is seriously mismatched with the actual tag population size. The performance will degrade further when IIoT is deployed in the error-prone channel environment. To optimize the identification performance of RFID system in an error-prone channel environment, we propose an efficient early frame breaking strategy based anticollision algorithm (EFB-ACA) with channel awareness. The EFB-ACA divides the whole tag identification process into two phases: convergence phase and identification phase. The function of convergence phase is to make the adjusted frame quickly converge to an appropriate size. The early frame breaking strategy is embedded in the convergence phase. Numerical results show that the proposed EFB-ACA algorithm outperforms the other methods on efficiency and stability in the error-prone channel. In addition, EFB-ACA algorithm also outperforms the other methods in the error-free channel.

1. Introduction

RFID is a key enabler of the Internet of Things (IoT), playing a crucial role in connecting low-/nonpowered devices to IoT environments. EPC C1 Gen2 [1] is the standard UHF RFID protocol devised to meet the demands and requirements of such applications. An RFID reader can communicate with hundreds of passive RFID tags within seconds, even at a distance of several meters away from the tags. The most remarkable virtue of the Gen2 standard is its lightweight and universality. Due to the shared nature of communication channel, a passive RFID system requires a collision arbitration protocol to serialize tag responses and mitigate collisions between the tag responses.

Recently, there have been three types of collision arbitration protocols to cope with tag identification, namely, Aloha-based [2–5], query-tree-based [6–8], and tree-splitting-based [9, 10] protocols. In Aloha-based protocols, the reader uses tag quantity estimation and frame size

adjustment strategies to identify tags. Specifically, the reader sends a query command containing a parameter F (F denotes the number of slots per frame) to allow the tags to return the IDs. The reader decodes the responses in each slot and distinguishes their states: collision, empty, and singleton. According to slot statistics in a frame, the reader can estimate the cardinality of unread tags and update the new frame for the next round. According to the analysis in previous works, the maximal system throughput of Aloha-based protocols is 0.368 [2–5]. In query-tree-based protocols [6–8], the reader queries the tags through probe commands. Each tag is required to be equipped with a prefix matching circuit, and it will respond only when the tag ID matches the prefix of the probe command. Once a collision is detected, the reader will update the query prefix according to the position of collision bits. Then, the reader uses the updated prefixes to interrogate the tags until all of them are successfully identified. In tree-splitting-based protocols, the reader continues to group the colliding tags with a

separation probability of 0.5, until a certain group contains only one tag.

In existing Aloha-based anticollision algorithms, tag population size estimation is a critical issue because it can help the reader to achieve the maximum efficiency when the frame size is the same as the tag population size. The authors in literature [11] utilized minimum distance to estimate the tag population size and adjusted the frame size based on the estimated tag cardinality. Schoute [12] concluded that the unread tag population size should be approximately 2.39 times of the number of collision slots. Chen in [13] proposed a feasible and easy-to-implement anticollision (FEIA) algorithm. In FEIA algorithm, when the first slot is an empty slot, the identification efficiency will be seriously constrained. The study in [14] firstly presented a subframe-based dynamic frame slotted Aloha (SUBF-DFSA) algorithm, which introduces a fixed early breaking scheme to tag identification process. However, the error-prone channel is not considered in above methods. Moreover, when the initial frame size is mismatched with the actual number of tags, the identification efficiency of these algorithms will be seriously affected.

Nowadays, RFID tags can be assembled almost anywhere. Considering environmental factors, RFID systems may not be able to maintain highly reliable communications [15]. The signal quality of RFID systems will be affected by the communication distance between the reader and tags [16]. The reader may not be able to successfully decode the weak signal returned by the tag; and it will cause the wrong statistical information of slot types in a frame [17]. Accordingly, the performance of anticollision algorithm will be severely weakened. Therefore, when designing an anticollision algorithm, the adaptability of the algorithm to the channel environment should be considered.

To cope with the above challenges, we propose an efficient early frame breaking strategy based anticollision algorithm (EFB-ACA) with channel awareness. The proposed EFB-ACA divides the tag identification process into two phases: convergence phase and identification phase; and various collision ratios are applied to identification phase to improve the identification efficiency. The core contributions of this paper can be summarized as follows:

- (1) An efficient and channel-aware anticollision algorithm is proposed. It can derive the actual number of singleton slots and collision slots in each frame based on the successful transmission possibility of identified tags, which indicates that EFB-ACA can be utilized in a variety of channel environments.
- (2) In our proposed EFB-ACA, the identification process can be divided into convergence phase and identification phase, where Bayesian estimation and various average collision ratios are utilized to improve the estimation accuracy of tag population sizes.
- (3) This paper reviews the most appropriate break point of the frame in detail. Our proposed approach makes full use of the early breaking strategy. It can not only

accelerate the convergence of frame sizes but also improve the stability of anticollision algorithm.

2. The Proposed Algorithm

In this section, we present the proposed algorithm in detail. Table 1 summarizes the notations used in the paper.

Due to the uncertainty of the channel, a singleton slot may be misjudged as a collision slot. Therefore, an error resilient method is required to improve channel adaptability of anticollision algorithms. EPC C1 Gen2 specifies the communication link between a reader and tags. The tags respond by reflecting the electromagnetic wave from the reader. Assume that the wavelength is λ and that the distance between the reader and the tag is d . The channel coefficients of the downlink and uplink are defined as h_d and h_u , respectively. Considering the large-scale fading, the downlink attenuation $|h_d|^2$ can be expressed as

$$|h_d|^2 = \left(\frac{\lambda}{4\pi d} \right)^2. \quad (1)$$

The uplink attenuation $|h_u|^2$ is equal to $|h_d|^2$ because the distances are equal. Let η denote the tag reflection loss. Assume that the transmitting power of the reader is defined as P_t , and the receiving power can be written as

$$P_r = P_t |h_d|^2 \eta |h_u|^2. \quad (2)$$

Let N denote the noise power, and the signal to interference plus noise ratio (SINR) can be calculated as

$$P_{\text{SINR}} = \frac{P_r}{N}. \quad (3)$$

Since the passive RFID system adopts PR-ASK modulation, the bit error ratio (BER) can be expressed as

$$R_{\text{BER}} = \frac{1}{2} \text{erfc} \left(\frac{\sqrt{R_{\text{SINR}}}}{2} \right), \quad (4)$$

where $\text{erfc}()$ is the complementary error function. Consider that the packet length is L , and the probability that a tag response can be successfully decoded by the reader can be expressed as

$$P_d = (1 - R_{\text{BER}})^L. \quad (5)$$

Accordingly, the reader can estimate the distance away from the tag based on the fixed transmitting power. When a tag is identified successfully, the reader can obtain the distance and calculate P_d . After identifying all tags, the reader can obtain the average successful probability \bar{P}_d . The actual number of singleton slots can be estimated as

$$S = \frac{S_d}{\bar{P}_d}. \quad (6)$$

Then, the actual number of collision slots can be estimated as

TABLE 1: Notations used in the paper.

Symbol	Description
F	Frame size
F_{sub}	Subframe size
N	Number of tags waiting to be identified
Nest	Estimated number of remaining tags
E	Number of empty slots in a frame
S	Number of actual singleton slots in a frame
C	Number of actual collision slots in a frame
S_d	Number of slots where the tag response is successfully decoded in a frame
C_d	Number of slots where the tag response fails to be decoded in a frame

$$C = C_d - \frac{(1 - \overline{P_d})S_d}{\overline{P_d}}. \quad (7)$$

Due to the estimation of channel quality, the proposed algorithm can improve robustness and identification efficiency over the previous work.

To accelerate convergence and improve estimation accuracy, early breaking strategy and Bayesian estimation are utilized in convergence phase. When the first m slots are collision slots, the actual tag population size is probably much larger than current frame size. The following slots are very likely to be collided. It can be viewed as current frame size is seriously mismatched with the actual tag population size. In this case, it is necessary to end the frame and adjust the frame size in advance. If early breaking strategy is used, the break point needs to be determined. The tag population size is a random variable N , which follows a discrete probability distribution. The number of tags in a subframe is also a random variable N_{sub} . Due to randomness of a tag to choose a slot, the relation between N_{sub} and N is

$$N_{\text{sub}} = \left(\frac{M_{\text{sub}}}{M}\right)N. \quad (8)$$

Then, the total number of tags N_E is written as

$$N_E = \left(\frac{M}{M_{\text{sub}}}\right)N_{\text{sub}}, \quad (9)$$

and $E(N_E)$ is the estimated tag population size, which is the expected value of N_E . $D(N_E)$ is the variance of N_E . The error of expected value can be calculated as

$$\text{Error}_E = \frac{|E(N_E) - E(N)|}{E(N)}. \quad (10)$$

The error of variance is

$$\text{Error}_D = \frac{|D(N_E) - D(N)|}{D(N)}. \quad (11)$$

Assume that the maximum error of variance is ε , and we can have

$$\text{Error}_D \leq \varepsilon. \quad (12)$$

Then, we can obtain

$$\sqrt{\frac{M^2 D(N_{\text{sub}})}{(1 + \varepsilon)D(N)}} \leq M_{\text{sub}} \leq \sqrt{\frac{M^2 D(N_{\text{sub}})}{(1 - \varepsilon)D(N)}}. \quad (13)$$

Therefore, given a probability distribution and ε , the break points can be calculated. As Figure 1 shows, in the tag identification process, the distribution of tag population size is close to normal distribution. To obtain appropriate break points, we choose the normal distribution $N(M, (M/16)^2)$ as the distribution of N and set $\varepsilon=0.03$. Then, the break points shown in Table 2 can be derived.

At the breaking point of a frame, the distribution of tag population size needs to be updated based on Bayesian estimation. The updating function is

$$P(N|E, S, C) = \frac{P(N)P(E, S, C|N)}{P(E, S, C)}, \quad (14)$$

where $P(N|E, S, C)$ is the updated distribution and $P(N)$ is the previous distribution. Conditional probability $P(E, S, C|N)$ can be calculated by

$$P(E, S, C|N) = \frac{M!}{E!S!C!}P(E)P(S|E)P(C|E, S). \quad (15)$$

where $P(E)$ is the probability of E empty slots; $P(S|E)$ and $P(C|E, S)$ are the conditional probability of S successful slots and C collision slots. $P(E, S, C)$ can be considered as a constant. It can be ignored during updating distribution. When the distribution is updated, normalization should be completed for the probability distribution. Since the identified tags are no longer involved in the following identification process, $P(N - S|E, S, C)$ will be the new probability distribution $P(N)$ of the random variable N in the next frame. The number of rest tags will be estimated as the expected value $E(N)$. We limit the maximum adjustment multiple of frame sizes as 8 for the consideration of stability.

Since the frame sizes are restricted to the power of 2, various tag population sizes may be matching with the same frame size. The corresponding relation between the number of tags and optimal frame size is shown in Table 3.

Once the frame size is matching with the estimated tag population size, the proposed algorithm will enter identification phase, where the frame size will almost match with the tag population size all the time. The aim of this phase is to improve identification efficiency. In such phase, the early breaking strategy is no longer utilized. For the sake of estimating the tag population size, it is needed to compute the number of tags in a collision slot. The average collision ratio R_c can be calculated by

$$R_c = \frac{1}{(N_{\text{max}} - N_{\text{min}} + 1)} \sum_{N=N_{\text{min}}}^{N_{\text{max}}} \frac{N - N_s}{M_c}. \quad (16)$$

where N_{max} and N_{min} are the maximum and minimum number of tags corresponding to the current frame size. N_s is the average successful tag population size, which is written as

$$N_s = N \left(1 - \frac{1}{M}\right)^{N-1}. \quad (17)$$

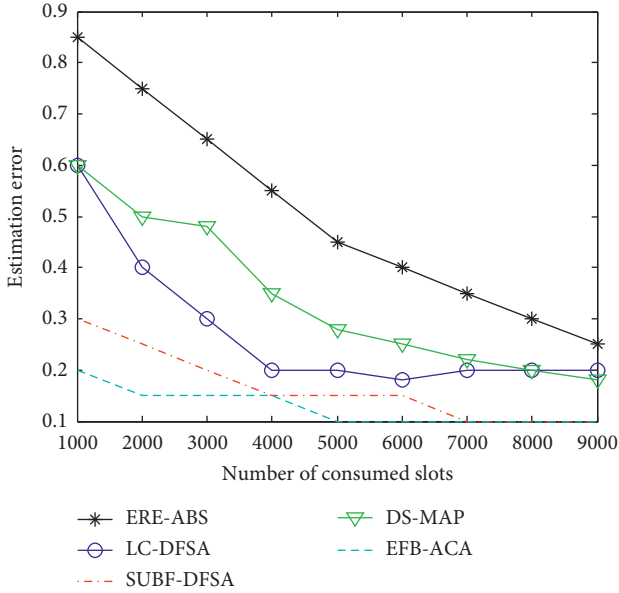


FIGURE 1: Performance comparison of the convergence phase in error-prone channel.

TABLE 2: Breaking point.

Frame size	Subframe size
16	16
32	16
64	16
128	16
256	16
512	32
1024	32
2048	32

TABLE 3: Optimal frame sizes for various tag population sizes.

Tag population size	Frame sizes
2–5	4
6–11	8
12–22	16
23–44	32
45–88	64
89–177	128
178–354	256
355–709	512
710–1419	1024

M_c is the expected collision slots, which can be obtained by

$$M_c = M - N_s - M \left(1 - \frac{1}{M}\right)^N. \quad (18)$$

Then, the number of tags in the collision slot can be derived. Therefore, the number of rest tags will be estimated as R_c times of the actual collision slot population size C . Then, the corresponding optimal size of the next frame will be confirmed.

Tag identification process enters the convergence phase first and transfers to identification phase when the frame size becomes matching with number of rest tags. In the convergence phase, early breaking method is utilized to accelerate the convergence process. At the check slot of each frame, whether current frame size is matching with the estimated tag population size or not will be examined. If not, current frame will break and update the distribution of the tag population size. Then, the reader launches a new round with an appropriate frame size. If the current frame size is matching with the estimated tag population size, the frame will continue and enter identification phase at the end of the frame.

In the identification phase, simple tag estimation is used to improve the efficiency and reduce the complexity. After each frame, the related information of the frame will be recorded. According to the current frame size and the number of collision slots, the number of rest tags will be accurately estimated. Then the optimal frame size can be determined. The above procedures will be repeated until there are no collisions.

3. Experimental Study

This paper evaluates the performance of our proposed EFB-ACA algorithm, SUBF-DFSA, ERE-ABS, DS-MAP, and LC-DFSA in the error-prone channel. To maintain the convergence, Monte Carlo method is used in the simulations [18–21]. Figure 1 compares the convergence phase duration of various algorithms in error-prone channel. Benefiting from the early breaking strategy and Bayesian estimation method, the proposed EFB-ACA can quickly pass the convergence phase and choose an appropriate frame size. Due to the fixed break points, the convergence speed of SUBF-DFSA is slower than that of the proposed algorithm. Because LC-DFSA and DS-MAP adjust the frame size only at the end of each frame, they will consume more slots to obtain an optimal frame size. As for ERE-ABS, it consumes most number of slots to enter into identification phase, resulting into the worst convergence speed among these algorithms.

Figure 2 shows the efficiency in the identification phase of various algorithms in the error-prone channel. The initial frame sizes of various algorithms are varied. The proposed algorithm estimates the number of rest tags by different collided ratios, which leads to the accurate estimation result. It has the highest identification efficiency among these algorithms. The other methods are affected by the error-prone channel. The efficiencies of ERE-ABS and LC-DFSA are relatively less affected by initial frame size; and SUBF-DFSA is relatively more affected by initial frame size. The reason is as follows. In an error-prone channel environment, an inappropriate frame size will cause a large difference between the collision probability and the empty probability, leading to inaccurate estimation results of tag cardinality.

Figure 3 presents the identification efficiency of different algorithms in the error-prone channel. We can find that the proposed algorithm can identify the same amount of tags with the least slots. It fully utilizes early break scheme and

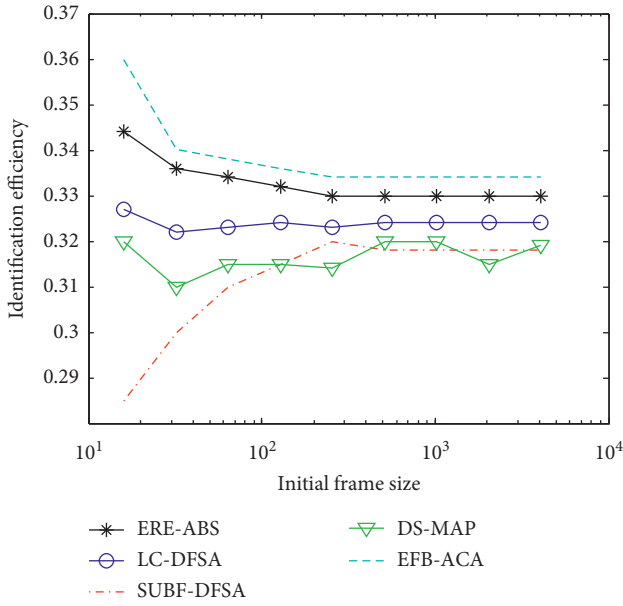


FIGURE 2: Performance comparison of the identification efficiency in various initial frame sizes.

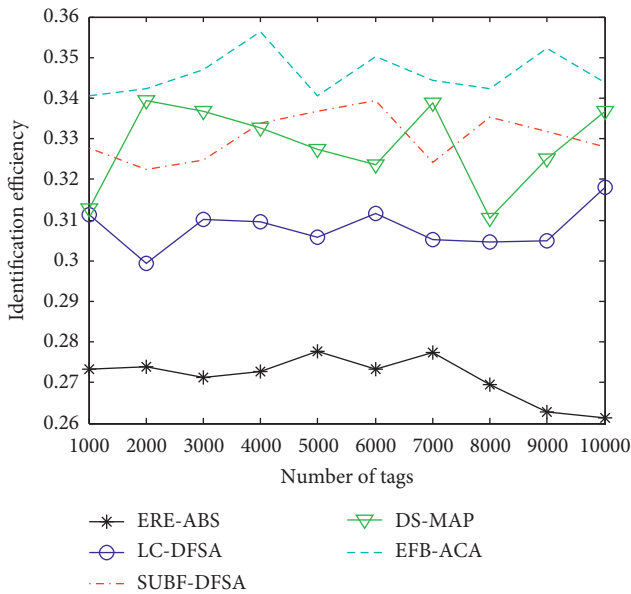


FIGURE 3: Comparison of identification efficiency in error-prone channel.

Bayesian estimation. Then, it achieves the highest identification efficiency in both convergence stage and identification stage. However, for the other reference methods, the inaccurate information of singleton and collision slots leads to the estimation accuracy deterioration. Then, the identification efficiency will be affected.

Figure 4 presents the effect of different initial frame sizes on the access efficiency in the error-prone channel. The number of tags is set as 1,000. As we can see from the results, the proposed algorithm can achieve stable and high access efficiency no matter the initial frame size is. The efficiency is

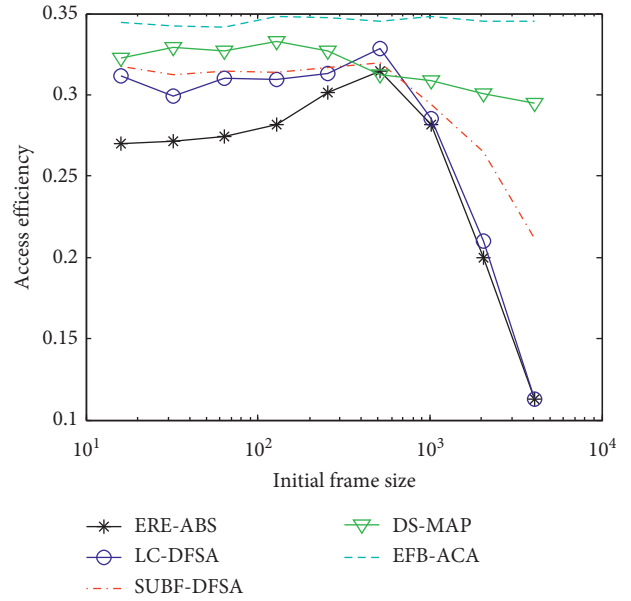


FIGURE 4: Influence of initial frame size in error-prone channel.

not affected by the initial frame size because of the reasonable early break scheme. When the initial frame size is small, SUBF-DFSA can realize stable efficiency. However, the stability will be worse as the initial frame size further increases due to its fixed early points. Because the early break method is not utilized in ERE-ABS and LC-DFSA, the initial frame size greatly affects the identification efficiency; and when the initial frame size is close to the number of tags, it can achieve higher identification efficiency.

4. Conclusion

In order to improve tag identification efficiency and channel adaptability of anticollision algorithms in RFID systems, a Bayesian estimation and early break based anticollision algorithm with channel awareness is proposed. We model the communication links between the reader and tags and carry out theoretical analysis of the proposed algorithm. The reader can estimate the link condition through the fixed transmitting power. It helps to obtain the real slot status; and the whole tag identification process is divided into two stages: convergence stage and identification stage. The main purpose of convergence stage is to make the frame size quickly converge to an appropriate value; and the tag identification is mainly performed in identification stage. Early break scheme and Bayesian estimation are applied to the convergence stage to accelerate convergence and improve the estimation accuracy. Different collided ratios are used in the identification stage, which can improve the identification efficiency. Through our simulations, we can find that the proposed algorithm outperforms the other algorithms on both stages in the error-prone channel; and the initial frame size has little effect on efficiency of the proposed algorithm. Besides, even in the error-free channel, the proposed algorithm can also outperform the other algorithms. The identification efficiency of the proposed

algorithm can reach 0.36, which is close to the theoretical optimal performance.

Data Availability

The authors derived the writing material from different journals as provided in the references. A MATLAB tool has been utilized to simulate the authors' concept.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by China Postdoctoral Science Foundation under Grant nos. 2020M681959 and 2020TQ0291.

References

- [1] GS1, *EPC™ Radio-Frequency Identity Protocols Generation-2 UHF RFID Standard*, GS1, Brussels, Belgium, Release 2.1, 2018.
- [2] J. Su, R. Xu, S. Yu, B. Wang, and J. Wang, "Idle slots skipped mechanism based tag identification algorithm with enhanced collision detection," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 5, pp. 2294–2309, 2020.
- [3] J. Su, Z. Sheng, A. X. Liu, Z. Fu, and Y. Chen, "A time and energy saving-based frame adjustment strategy (TES-FAS) tag identification algorithm for UHF RFID systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 2974–2986, 2020.
- [4] H. Chen, K. Liu, C. Ma, Y. Han, and J. Su, "A novel time-aware frame adjustment strategy for RFID anti-collision," *Computers, Materials & Continua*, vol. 57, no. 2, pp. 195–204, 2018.
- [5] W.-T. Chen, "Optimal frame length analysis and an efficient anti-collision algorithm with early adjustment of frame length for RFID systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3342–3348, 2016.
- [6] X. Jia, Q. Feng, and L. Yu, "Stability analysis of an efficient anticollision protocol for RFID tag identification," *IEEE Transactions on Communications*, vol. 60, no. 8, pp. 2285–2294, 2012.
- [7] J. Su, Y. Chen, Z. Sheng, Z. Huang, and A. X. Liu, "From M-ary query to bit query: a new strategy for efficient large-scale RFID identification," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2381–2393, 2020.
- [8] H. Landaluce, A. Perillos, E. Onieva, L. Arjona, and L. Bengtsson, "An energy and identification time decreasing procedure for memoryless RFID tag anticollision protocols," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4234–4247, 2016.
- [9] J. Su, Z. Sheng, L. Xie, G. Li, and A. X. Liu, "Fast splitting-based tag identification algorithm for anti-collision in UHF RFID system," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2527–2538, 2019.
- [10] J. Su, Z. Sheng, A. X. Liu, Y. Han, and Y. Chen, "A group-based binary splitting algorithm for UHF RFID anti-collision systems," *IEEE Transactions on Communications*, vol. 68, no. 2, pp. 998–1012, 2020.
- [11] H. Vogt, "Multiple object identification with passive RFID tags," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 6–13, Yasmine Hammet, Tunisia, October 2002.
- [12] F. Schoute, "Dynamic frame length ALOHA," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 565–568, 1983.
- [13] W.-T. Chen, "A feasible and easy-to-implement anticollision algorithm for the EPC global UHF class-1 generation-2 RFID protocol," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 2, pp. 485–491, 2014.
- [14] J. Su, Z. Sheng, D. Hong, and V. C. M. Leung, "An efficient sub-frame based tag identification algorithm for UHF RFID systems," in *Proceedings of the 2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, Kuala Lumpur, Malaysia, May 2016.
- [15] A. Bekkali, S. Zou, A. Kadri, M. Crisp, and R. Penty, "Performance analysis of passive UHF RFID systems under cascaded fading channels and interference effects," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1421–1433, 2015.
- [16] D. Benedetti, G. Maselli, C. Petrioli, and M. Piva, "The impact of external interference on RFID anti-collision protocols," *IEEE Networking Letters*, vol. 1, no. 2, pp. 76–79, 2019.
- [17] J. Park and T. Lee, "Error resilient estimation and adaptive binary selection for fast and reliable identification of RFID tags in error-prone channel," *IEEE Transactions on Mobile Computing*, vol. 11, no. 6, pp. 959–969, 2012.
- [18] I. Filho, I. Silva, and C. Viegas, "An effective extension of anti-collision protocol for RFID in the industrial Internet of Things (IIoT)," *Sensors*, vol. 18, no. 12, p. 4426, 2018.
- [19] Y. Chen, J. Su, and W. Yi, "An efficient and easy-to-implement tag identification algorithm for UHF RFID systems," *IEEE Communications Letters*, vol. 21, no. 7, pp. 1509–1512, 2017.
- [20] J. Su, X. Zhao, D. Hong, Z. Luo, and H. Chen, "Q-value fine-grained adjustment based RFID anti-collision algorithm," *IEICE Transactions on Communications*, vol. E99.B, no. 7, pp. 1593–1598, 2016.
- [21] J. Su, Z. Sheng, A. X. Liu, Y. Han, and Y. Chen, "Capture-aware identification of mobile RFID tags with unreliable channels," *IEEE Transactions on Mobile Computing*, p. 1. In press, 2020.

Research Article

The Impact of Urban Sprawl and Smart City Construction on Regional Coordination

Liangfeng Hao ¹, Xue Chen ², and Chengpeng Min ³

¹School of Business, Suzhou University of Science and Technology, Suzhou 215009, China

²School of Business, Moutai Institute, Renhuai 564507, China

³Guizhou Urbanization Development Center, Guiyang 550003, China

Correspondence should be addressed to Xue Chen; xue.chen@mtxy.edu.cn

Received 12 January 2021; Revised 31 March 2021; Accepted 10 May 2021; Published 19 May 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Liangfeng Hao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The construction of smart cities has promoted the process of urbanization and sustainable urban sprawl, which may accelerate regional coordination by enhancing the spatial correlation among the cities. Firstly, this paper built the mechanism for the impact of urban sprawl and smart city construction on regional coordination and adopted the corrected night light data as the index of economic measurement, using the dynamic fixed effect spatial Dubin model to test theoretical mechanism. It is found that urban sprawl has strongly promoted the regional coordination, which is especially obvious among the neighboring cities. However, the construction of smart cities is not conducive to regional coordination, only when interacting with urban sprawl. The results of robustness check and endogenous treatment are consistent with the baseline regression. Further research shows that urban sprawl restricts the positive effect of industrial agglomeration but could promote economic growth and regional coordination through smart city construction. The policy enlightenment lies in that smart city construction should be promoted, so as to improve economic growth, and smart city network and urban sprawl should be synchronously promoted to accelerate regional coordination.

1. Introduction

China's economic development has achieved remarkable achievements in the process of reform and opening up. However, with the inflow of production factors to the east, the difference among eastern coastal areas and inland areas is gradually widening. More seriously, the location advantage of the eastern area is still potential driving force for attracting production factors (Liu [1]). Meanwhile, the level of urbanization has been rapidly improved, which has led to the continuous expansion of the urban space, experiencing rapid, discontinuous and low-density expansion (Liu et al. [2]). The research shows that China's urban space increased by 43.5% from the year of 2000 to 2010, and the land expansion was significantly faster than the improvement of population, leading to the typical characteristics of increasing urban population and decreasing population density (Jiang and Xi [3]). Figure 1 shows the mutative trend

of urbanization rate and regional difference from 1978 to 2018. It could be found that the urbanization level increased rapidly after 1998, while regional differences tended to be flat around 1998, which shows obvious downtrend around (2004).

This paper focuses on the relationship between urban sprawl and regional coordinated development based on the impact of smart city construction. Smart cities are formed by the continuous development and evolution of digital cities. It is an inevitable stage for the level of urbanization rising to a certain degree, which has become an important driving force for the development of urbanization in China (Lv et al. [4]). The path of smart city construction is to improve human resources, information infrastructure, and urban innovation capabilities. Through the widespread use of smart application platforms, the government's urban management capabilities and service levels could be improved, which could connect different regions and promote the spread of urban

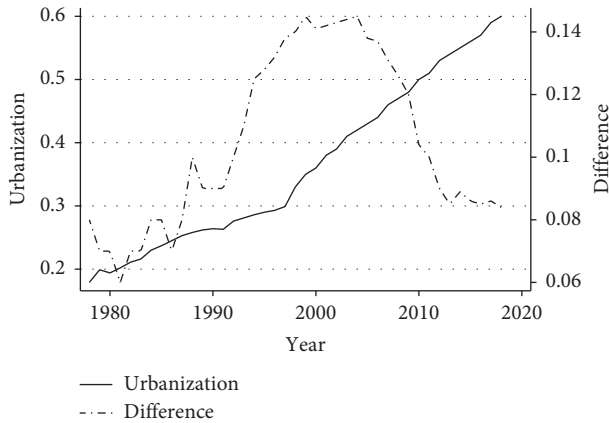


FIGURE 1: The evolution trend of urbanization rate and regional differences from 1978 to 2018.

space (Lu and Yang [5]). We try to explore how smart city construction affects the process of urban sprawl, and the impact of urban sprawl and smart city construction on regional coordination.

The remainder of the study is organized as follows. Firstly, we intend to review the relevant literature on smart city, urban sprawl, and regional coordination systematically. Secondly, we explain the mechanism of how smart city construction affects urban sprawl, which might promote regional coordination. Thirdly, based on the process of modifying the urban night light data, we intend to measure the level of urban sprawl combined with LandScan population dynamics statistics. Fourthly, we intend to use spatial Durbin dynamic panel model with fixed effect for empirical analysis, and urban surface roughness is used for endogenous treatment. Finally, robustness test of the empirical results is conducted, and we try to interpret the mechanism between urban sprawl and regional coordination by mediating effect test through dynamic industrial agglomeration.

2. Literature Review

With the rapid development of new generation of information technology, the concept of smart city proposed by IBM has become new idea for urban development. Smart city construction has become an important driving force of urbanization in China. It is required not only to solve economic problems in the process of urban development, but also to improve social, cultural, and environmental issues in urban governance. Domestic and foreign research on the evaluation index system of smart city construction mainly focuses on the dimensions of network and other infrastructure, economic development and industrial adjustment, social management and services, and value guidance and realization (Xu et al. [6]). Chen et al. [7] found that smart city is formed by key elements such as organization, technology, governance, policy environment, communities, economy, infrastructure, and natural environment. Smart city construction is required to provide efficient governance and service systems, upgraded

industrial structure, and comfortable and superior living environment (Chourabi et al. [8]; Alawadhi et al. [9]).

The phenomenon that goes hand-in-hand with the construction of smart cities is urban sprawl. Urban sprawl is featured by dynamic evolution of urban spatial structure, manifested by the feature of geographical space expansion and population density decrease, making the urban structure more decentralized and polycentric (Glaeser and Khan [10]). Liu et al. [2] explained the causes of urban sprawl from the perspective of market uncertainty. Due to spatial dispersion of elements and decreasing level of industrial agglomeration, urban sprawl may incur efficiency losses (Lee and Gordon, [11]). The literatures are also more inclined to support the conclusion that urban sprawl may restrict the process of production efficiency improvement or economic growth. The mechanism lies in the decline of industrial agglomeration caused by sprawl, which intends to increase the cost of “face-to-face” communication (Fallah et al. [12]; Farber and Li [13]; Di Liddo [14]; Qin and Liu [15]). Therefore, the positive effect of industrial agglomeration is weakened. However, for cities with too large scale, which may lead to uneconomic agglomeration, urban sprawl could reduce the production cost and housing cost in turn, which is conducive to economic or productivity improvement. In addition, some studies focused on public welfare, and ecological environment under the influence of urban sprawl (Sun and Wan [16]; Jiang [17]).

Since urban sprawl and smart city construction could change the spatial distribution of elements and accelerate the flow of elements, it could be inferred that the process attaches impact on regional coordination. The literatures of research on regional coordination have accumulated plentiful achievement, among which different results are presented. In the early studies, it is intended to believe that there is no convergence throughout the whole country, showing conditional convergence, or club convergence (Lin and Liu [18]; Xu and Li [19]). In recent years, it is more likely to agree that there is stronger convergence across the country. The characteristics of club convergence are prominent due to the development of urban agglomerations (Tan et al. [20]; He and Chen [21]). However, few studies believe that there is no obvious convergence in China, among which expansion of differences exists (Shi and Ren [22]). With the application and development of spatial methods, most studies need to consider spatial spillover effects among the regions. It is found that the regions showed stronger convergence trend under the condition of spatial effect. Ignoring spatial effect may even lead to diametrically opposite conclusions. In addition, the research on regional productivity synergy emerged in the latest years, with the increasing contribution of total factor productivity on economic growth (Yu [23]).

Through literature review, we find that smart city construction is the type of upgraded urbanization, and how could urban sprawl and smart city construction affect regional coordinated development needs more evidence and further explanation. The following parts focus on the mechanism about the impact of urban sprawl and smart city construction on regional economic coordination, and the empirical test based on the city-level data.

3. Mechanism Model

The theoretical model on the impact of urban sprawl on economic convergence is built based on the neoclassical growth model, which is expanded by the model of Mankiw et al. [24] and López et al. [25]. Firstly, the simple economic growth model could be expressed as

$$Y(it) = A(it)K(it)^{\tau_k}H(it)^{\tau_h}L(it)^{1-\tau_k-\tau_h}, \quad (1)$$

in which τ_k and τ_h indicate the share of physical capital and human capital, respectively. There is spatial spillover effect on the economic growth of cities and neighboring spatial units, and there are centripetal force and centrifugal force featured by nonlinear change among the cities, which could affect the spatial balance. We assume that technology $A(it)$ with shared characteristics depends on the technology level of the cities and neighboring areas, and as the scale of neighboring cities expands, the spillover effect presents a nonlinear change among the cities. It could be expressed as

$$A(it) = \Delta_{it} \left(k_{pit}^{\tau_k} h_{pit}^{\tau_h} \right)^{s^y}, \quad (2)$$

in which Δ_{it} indicates the exogenous parameters, and k_{pit} and h_{pit} indicate the unit physical capital and human capital of adjacent spatial areas, respectively. s^y indicates the spatial spillover effect, in which s represents the level of urban sprawl, and y indicates the level of urban intelligence. Both sizes are between (0, 1). We can derive model (3) by putting (2) into (1), which is expressed as

$$y(it) = \Delta_{it} k(it)^{\tau_k} h(it)^{\tau_h} \left(k_{pit}^{\tau_k} h_{pit}^{\tau_h} \right)^{s^y}. \quad (3)$$

According to the equilibrium of the neoclassical growth model, we derive the following formulas (4) and (5):

$$g_k = \frac{\dot{k}}{k} = s_k k^{-(1-\tau_k)} h^{\tau_h} k_p^{s^y \tau_k} h_p^{s^y \tau_h} - (n + g + \delta), \quad (4)$$

$$g_h = \frac{\dot{h}}{h} = s_h k^{\tau_k} h^{-(1-\tau_h)} k_p^{s^y \tau_k} h_p^{s^y \tau_h} - (n + g + \delta), \quad (5)$$

in which n indicates the population growth rate, and g indicates technological progress. s_k and s_h indicate the output share of physical capital and human capital, respectively, and δ indicates the capital depreciation rate. When $\dot{k}(t)$ and $\dot{h}(t)$ are equal to zero, and the hypothesis of $\tau_k + \tau_h < 1$ is satisfied simultaneously, the steady-state capital stock k^* and h^* could be deduced as

$$k^* = \left(\frac{s_k^{1-\tau_k} s_h^{\tau_h} k_p^{s^y(1-\tau_k)} h_p^{s^y \tau_h}}{n + g + \delta} \right)^{1/(1-\tau_k-\tau_h)}, \quad (6)$$

$$h^* = \left(\frac{s_k^{\tau_k} s_h^{(1-\tau_k)} k_p^{s^y \tau_k} h_p^{s^y(1-\tau_k)}}{n + g + \delta} \right)^{1/(1-\tau_k-\tau_h)},$$

and formula (7) could be derived, combining with formula (3):

$$y^* = \left(\frac{s_k^{\tau_k} s_h^{\tau_h} k_p^{s^y \tau_k} h_p^{s^y \tau_h}}{(n + g + \delta)^{\tau_k + \tau_h}} \right)^{1/(1-\tau_k-\tau_h)}. \quad (7)$$

It is derived that the per capita output that reaches the long-term equilibrium state is related not only to the share of capital gains, but also to the share of capital in neighboring units and s^y which is used to measure the impact of the size of neighboring cities. According to the neoclassical growth model, the relationship between actual average output and steady-state output could be expressed as

$$\frac{d \ln(y(it))}{dt} = \lambda [\ln(y^*(it)) - \ln(y(it))], \quad (8)$$

in which $\lambda = (n + g + \delta)(1 - \tau_k - \tau_h)$.

The economic growth equation could be expressed as formula (9), which is Taylor expansion of formula (8):

$$\ln(y(it)) - \ln(y(0)) = \varepsilon - \frac{(1 - e^{-\lambda t})s^y}{1 - \tau_k - \tau_h} \ln y(0) + \frac{(1 - e^{-\lambda t})s^y}{1 - \tau_k - \tau_h} y_p(0) + s^y g_{y_p} + \frac{(1 - e^{-\lambda t})}{1 - \tau_k - \tau_h} [\tau_k (\ln s_k + \ln(n + g + \delta))] + \tau_h (\ln s_h + \ln(n + g + \delta)), \quad (9)$$

in which $\varepsilon = (1 + s^y)g - (1 - e^{-\beta t})(1 - s^y/1 - \tau_k - \tau_h)$ ($\ln \Delta_0 + gt$).

According to $(1 - e^{-\lambda t})s^y/1 - \tau_k - \tau_h$ in formula (9), we could draw the proposition that the convergence trend could be strengthened together by urban sprawl and the smart city construction.

4. Data, Indicator, and Modeling

4.1. Data. Gross domestic product (GDP) is an important indicator that is used to measure economic development or growth. However, the biggest limitation lies in the statistical comparison among the regions. In addition, local officials

may have strong incentives to falsify data under China's administrative assessment system linked to GDP. Thus, some research began to use the global night light data released by the National Oceanic and Atmospheric Administration (NOAA), in order to achieve the comparability and reflect the level of economic development (Henderson et al. [26]). NOAA has published data from 1992 to 2013, including the satellites of F10, F12, F14, F15, F16, and F18, forming a total of 34 images, as well as thirty-four data images. The spatial light resolution is 30 seconds, and the gray value interval is (0, 63). The data needs to be corrected since the original night lights are not stable, and there are two sets of satellite observation data in some years. We take the average value when there are different satellite observations, and we uniformly set the value as zero when the satellite DN observation value is zero in any year, attributed to the abnormal fluctuation of light.

Due to the large instability of night light data, there are two sets of light data in individual years. The data needs to be corrected concerning the aging and switching of observing satellites may reduce the comparability of data in different years.

4.1.1. Baseline Correction. Since the light data is observed by multiple groups of satellites, it is necessary to set reference point for mutual correction. We take the city of Hegang in Heilongjiang province as the reference area and take the data of F18 satellite as the benchmark. The correction model is expressed as

$$SP_i = 0.5 * (LP_i - HP_i) + 0.5. \quad (10)$$

The values of a , b , and c could be simulated by equation (10), in which $DN_{HG, F18-2013}$ represents the Hegang grid value observed by the F18 satellite in 2013, while $DN_{HG, F-T}$ indicates Hegang grid values observed by different satellites in other years. Through the equation, thirty-three estimated values of a , b , and c could be obtained, to bring the estimated values a , b , and c into function (11) composed by the independent variables of the F18 observations of different cities in 2013, so as to obtain the corrected light values.

4.1.2. Outlier Value Handling. Since there are two sets of satellite observations in specific year, it could be treated by taking averaging value. In addition, the value is set as zero if the DN value is zero, even if another group of observing satellites shows nonzero value. The reason lies in that another set of nonzero value might be caused by abnormal fluctuations.

4.2. Indicator

4.2.1. Urban Sprawl Indicator. Referring to the method of Fallah et al. [27], the index for measuring urban sprawl is set as follows:

$$SP_i = 0.5 * (LP_i - HP_i) + 0.5, \quad (11)$$

in which LP_i is the ratio of population density, which is lower than the national average density, while HP_i represents the ratio of the population density that is higher than the national average density. However, it is necessary to concern land expansion since urban sprawl is reflected not only in population expansion, especially under the background that the land urbanization is faster than population urbanization in China. Thus, we further modify the urban sprawl index as follows, referring to Qin and Liu [15]:

$$SA_i = 0.5 * (LA_i - HA_i) + 0.5, \quad (12)$$

in which LA_i is the ratio of land density, which is lower than the national average density, while HA_i represents the ratio of the land density that is higher than the national average density. The new indicators for urban sprawl could be further synthesized based on equations (11) and (12):

$$Sprawl_i = \sqrt{SP_i * SA_i}. \quad (13)$$

The urban sprawl value interval constructed by equation (13) is (0, 1), and the closer the value to 1, the higher the sprawl level. LandScan global population dynamics statistics provide valuable dataset for obtaining the population distribution since there is no detailed domestic population data in the cities. In addition, we integrated the two databases of night light data and LandScan global population dynamics data, so as to extract grids with DN value greater than 10 and population density greater than 1000 people/km².

4.2.2. Dynamic Industrial Agglomeration. We incorporate the urban geographic area into the construction of agglomeration index, in order to reflect the impact of urban sprawl. In addition, the level of industrial agglomeration should also be reflected in the process of firms' entering or exiting, since the research is usually limited to the stock level when constructing industrial agglomeration indicators. The industrial agglomeration index is built as follows.

Step 1. Taking the firms f in different industries added up to the city-industry level from three dimensions of labor, capital, and output and then divided by geographic area so as to obtain the spatial intensity in city-industry dimensions, the formula could be formed as

$$\begin{aligned} de_{cit} &= \sum_f \frac{\text{employment}_{fcit}}{\text{area}_{ct}}, \\ da_{cit} &= \sum_f \frac{\text{asset}_{fcit}}{\text{area}_{ct}}, \\ do_{cit} &= \sum_f \frac{\text{output}_{fcit}}{\text{area}_{ct}}, \end{aligned} \quad (14)$$

in which c represents the city, and i is the urban industry. de_{cit} , da_{cit} , do_{cit} indicate employment, asset, and output, respectively.

Step 2. To integrate the process of firms' entry and exit into the index, we use $(1 + p_{cit})$ as the weight to add up the spatial intensity of the previous year, in which p_{cit} is the net increase share of industry i . The formula could be formed as

$$\begin{aligned} ade_{cit} &= (1 + p_{citi}) * de_{cit-1}, \\ ada_{cit} &= (1 + p_{cita}) * da_{cit-1}, \quad \text{and,} \\ ado_{cit} &= (1 + p_{cito}) * do_{cit-1}, \end{aligned} \quad (15)$$

in which p_{citi} , p_{cita} and p_{cito} indicate the net increase share of employment, asset, and output, respectively.

Step 3. To calculate the share of industry i in city c in different dimensions, the final industrial dynamic agglomeration index in three dimensions could be obtained as

$$\begin{aligned} \text{density_adl}_{ct} &= \sum_i ada_{cit} * sl_{cit}, \\ \text{density_ada}_{ct} &= \sum_i ada_{cit} * sa_{cit}, \quad \text{and,} \\ \text{density_ado}_{ct} &= \sum_i ado_{cit} * so_{cit}, \end{aligned} \quad (16)$$

in which sl_{cit} , sa_{cit} and so_{cit} indicate the total share of employment, asset, and output, respectively.

In addition to the database of night light data and LandScan global population dynamics data, China industrial enterprise database and China City Statistical Yearbook are used to calculate the indicator.

4.3. Spatial Estimation Model. There are spatial correlations among neighboring regions according to the "First Law of Geography," and ignoring spatial factors may lead to deviations in empirical results. We use Moran index to test the spatial correlation of cities, and Table 1 gives the results of estimation. It could be found that there is positive spatial correlation among the cities from 2000 to 2013.

In general, we intend to measure regional coordination from the perspective of spatial convergence. We rejected the hypothesis that the spatial Durbin model could be simplified to spatial lag model or spatial error model by LR test. Therefore, the spatial Durbin fixed effects model is used as the baseline regression equation combined with Hausmann test. The equation is formed as

$$\text{Lnlt}_{it} - \text{Lnlt}_{it-1} = \lambda W_n (\text{Lnlt}_{it} - \text{Lnlt}_{it-1}) + \gamma \text{Lnlt}_{it-1} + \beta \text{Lnspawl}_{it} + \theta W_n * \text{Lnspawl}_{it} + \gamma \text{LnX}_{it} + \eta W_n \text{LnX}_{it} + \mu_i + \zeta_t + \xi_{it}, \quad (17)$$

in which Lnlt_{it} is indicated as night light data, Lnspawl_{it} is used to represent urban sprawl, and X_{it} are other control variables. μ_i and ζ_t represent the fixed effects of cities and time, and ξ_{it} is random error term. Particularly, W_n could be separated into three dimensions of weight matrix, including space weight matrix of adjacency space, physical distance, and economic distance. The construction of the adjacent space weight matrix is related to whether the cities are adjacent to each other. If two cities are adjacent to each other, the weight is assigned to 1, or it is 0. Although the adjacent space weight matrix has been widely used in spatial econometric analysis, the assumption that there is no correlation among nonadjacent cities is rather not reasonable. Therefore, this paper takes the reciprocal of the shortest road distance between the two cities as the weight of the spatial matrix, so as to further measure the location relationship among the cities. The formula could be written as $w_{ij}^d = 1/d_{ij}$. It is worth noting that the space weight matrix of physical distance ignored the economic connection among the cities. The spatial connection among the cities is not only related to distance or whether they are adjacent, but also related to the economic ties. The space weight matrix of economic distance could be written as

$$w_{ij}^e = w_{ij}^d \text{diag}\left(\frac{\bar{Y}_1}{\bar{Y}}, \frac{\bar{Y}_2}{\bar{Y}}, \dots, \frac{\bar{Y}_n}{\bar{Y}}\right), \quad (18)$$

in which w_{ij}^d is the space weight matrix of physical distance, and \bar{Y}_i is the average value of city lights, which could be written as $\bar{Y}_i = 1/(T_{\max} - T_{\min} + 1) \sum_{T_{\min}}^{T_{\max}} Y_i$. \bar{Y} is the unique average value of city lights, which could be written as $\bar{Y} = 1/n(T_{\max} - T_{\min} + 1) \sum_{i=1}^n \sum_{T_{\min}}^{T_{\max}} Y_i$.

X_{it} indicates the other control variables that affect economic coordination, such as savings rate. We take the compound expression of $(n + g + d)$ as the control variable according to (Yu [23]), in which n represents the population growth rate, g represents technological progress rate, and d represents depreciation rate. The calculation method of technological progress rate and depreciation rate is reported in Appendix. In addition, according to Xie and Wang [28], the construction of road infrastructure could influence the spatial distribution of firms, and the interdistrict transportation infrastructure connecting cities could promote the flow of production factors among the regions, thus promoting the coordinated development of regions. So, we take paved road area per capita to measure the level of road infrastructure. Meanwhile, due to the research of Yuan and Zhu [29], smart city is evaluated from three dimensions of information, human capital, and urban innovation activity. Limited to data availability, we take telecom business income as the index of cities' promotion of information technology and take the number of college students per ten thousand population to evaluate the level of human capital. In ad-

TABLE 1: Spatial correlation test.

year	Moran's I value	year	Moran's I value
2000	0.312*** (6.021)	2007	0.232*** (8.066)
2001	0.211*** (8.921)	2008	0.215** (2.014)
2002	0.212*** (7.321)	2009	0.202*** (5.054)
2003	0.243*** (2.832)	2010	0.205*** (2.811)
2004	0.258*** (6.031)	2011	0.208*** (5.587)
2005	0.206*** (7.067)	2012	0.366*** (9.054)
2006	0.218*** (7.954)	2013	0.404*** (9.065)

***, **, and * indicate significance at the 1%, 5%, and 10% level. Z values are reported in parentheses.

dition, Kou and Liu released the report of China's urban and industrial innovation, so that we could adopt the data to measure the level of urban innovation. The interactive items of three indicators are built to evaluate the construction of smart cities. Table 2 shows the statistical description of related variables.

5. Empirical Test

5.1. Baseline Regression Results. Table 3 presents the baseline results on how urban sprawl affects economic coordination, where Lnlt_{it-1} (Impact) is used to indicate the variable of Lnlt_{it-1} after introducing control variables. The result shows that the value of the coefficient Lnlt_{it-1} (Impact) increased significantly after introducing the variable of sprawl, indicating that the process of urban sprawl is beneficial to regional coordination. In addition, the value of Lnlt_{it-1} coefficient is the largest corresponding to the adjacent space weight matrix, indicating that the economic synergy among adjacent cities is stronger. Furthermore, the value of the coefficient Lnlt_{it-1} corresponding to the adjacent space weight matrix showed most obvious change under the impact of urban sprawl, reflecting more significant convergence among neighboring cities. However, the corresponding $W * \text{Lnsprawl}_{it}$ coefficient is significantly negative, indicating that the spatial spread of neighboring cities is not conducive to their own growth, which shows vicious competitive relationship.

5.2. Regression Results Based on Smart City Construction. Table 4 shows the regression result based on smart city construction. We adopt the variable of Lnsmart_{it} to indicate smart city construction, which is the interactive items of Lninformation , Lnlabor , and Lninnovation . In addition, the result is reported under the condition of space weight matrix of economic distance. Different from the baseline result, the absolute value of coefficient Lnlt_{it-1} shows decreasing trend after considering the influence of smart city construction, indicating that smart city construction is not conducive to promotion of economic coordination. However, smart city construction could promote regional economic growth due to the positive coefficient of Lnsmart_{it} . The interaction terms of urban sprawl and smart city construction are introduced to test the interactive effect. The result shows that the absolute value of coefficient Lnlt_{it-1} has been greatly improved after introducing the interactive item. In addition, the negative coefficient of Lnsprawl_{it} and positive coefficient of

$\text{Lninteraction}_{it}$ indicate that urban sprawl is not conducive to regional economic growth but shows positive effect through the interaction with smart city construction. The comparison among the results suggests that the conditions of interaction between urban sprawl and smart city construction could promote regional coordination, and the process of smart city construction could reduce the negative effect on economic growth, which is consistent with theoretical mechanism. Moreover, vicious competitive relationship still exists due to the negative value of $W * \text{Lninteraction}_{it}$.

5.3. Robustness Test. Suburbanization index is used to test the robustness of empirical results according to Liu et al. [2]. The method of suburbanization index is measured by the proportion of the population, which is three kilometers away from downtown. This paper takes night light data to identify the city center and delimits the three-kilometer range. Meanwhile, Landscan dataset is used to extract areas that include the lighting threshold and population density at the same time, so as to measure the level of suburbanization. Table 5 presents the robustness test, which shows no different conclusion from baseline result. The result suggests that the suburbanization could promote regional coordination. What needs to be emphasized is the negative effect on economic growth, which is more significant for suburbanization. The reason lies in that the calculation method of suburbanization emphasizes the expansion of land or population, especially excluding areas within three kilometers of the city center. However, smart city construction reduced the negative effect on economic growth and could impact positive effect on regional coordination.

5.4. Endogenous Processing. Although the spatial model could reduce the endogenous error by introducing spatial weight matrix, the generalized spatial two-stage least square method is used to deal with the endogenous term in order to avoid the deviation caused by the possible mutual influence between urban sprawl and economic growth. According to Burchfield et al. [30], in the areas with more uneven terrain, the difficulty of spatial expansion is increasing. Therefore, we use the reciprocal of urban surface roughness as an instrumental variable for urban sprawl. The surface roughness is obtained from the national urban digital elevation data, and the urban surface raster is extracted by Arc GIS software. The urban surface roughness could be obtained by

TABLE 2: Statistical description of related variables.

Variables	Sample size	Mean value	Minimum value	Maximum value	Standard deviation
Lnlt	3696	1.669	-1.790	4.781	1.172
Lnsprawl	3696	-0.856	-1.649	-0.268	0.219
Lnagg	3696	5.360	0.018	10.450	1.358
Lnngd	3696	-1.890	-7.142	-0.168	0.735
Lns	3696	-0.595	-1.115	-0.199	0.215
Lnroad	3696	1.961	-1.966	4.686	0.642
Lninformation	3696	12.289	10.044	15.554	0.845
Lnlabor	3696	-4.017	-6.986	-1.476	1.059
Lninnovation	3696	-0.733	-4.605	6.503	1.811

TABLE 3: Effect of urban sprawl on economic coordination.

Matrix classification	OLS	Adjacent space weight matrix	Space weight matrix of physical distance	Space weight matrix of economic distance
$Lnlt_{it-1}$	-0.014*** (-4.31)	-0.072*** (-8.60)	-0.067*** (-7.62)	-0.063*** (-7.59)
$Lnlt_{it-1}$ (impact)	-0.013*** (-3.85)	-0.076*** (-7.96)	-0.070*** (-5.79)	-0.065*** (-7.27)
$Lnsprawl_{it}$	0.023*** (6.42)	0.012* (1.88)	0.008 (1.19)	0.002 (0.34)
Lns_{it}	0.013 (1.58)	0.002** (2.52)	0.002** (2.54)	0.002*** (2.91)
$Ln(n+g+d)_{it}$	0.004*** (4.09)	0.019** (2.25)	-0.009 (-0.95)	0.004 (0.72)
$W * Lnsprawl_{it}$		-0.022** (-2.04)	-0.014 (-1.13)	0.046** (2.48)
$W * Lns_{it}$		0.002 (1.51)	0.002 (1.19)	-0.001 (-0.26)
$W * Ln(n+g+d)_{it}$		-0.021** (-2.07)	0.013 (1.04)	0.084*** (5.21)
Fixed effects	Yes	Yes	Yes	Yes
λ		0.553*** (24.59)	0.604*** (26.28)	0.145*** (3.67)
R^2	0.229	0.706	0.691	0.611
Observations	3696	3696	3696	3696

***, **, and * indicate significance at the 1%, 5%, and 10% level. Z values are reported in parentheses.

TABLE 4: Effect of smart city construction on economic coordination.

Variable	Impact of smart city	Impact of urban sprawl	Impact of interaction
$Lnlt_{it-1}$	-0.063*** (-7.59)	-0.063*** (-7.59)	-0.063*** (-7.59)
$Lnlt_{it-1}$ (impact)	-0.062*** (-7.56)	-0.065*** (-7.27)	-0.066*** (-7.38)
$Lnsmart_{it}$	0.001* (1.78)		0.002 (1.54)
$Lnsprawl_{it}$		-0.002 (-0.34)	-0.004* (-1.91)
$Lninteraction_{it}$			0.0006* (1.72)
$W * Lnsmart_{it}$	-0.0007** (-2.31)		-0.001* (-1.68)
$W * Lnsprawl_{it}$		0.046** (2.48)	-0.015 (-0.66)
$W * Lninteraction_{it}$			-0.0002*** (-3.21)
Control variable	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Λ	0.156*** (4.56)	0.145*** (3.67)	0.136*** (2.67)
R^2	0.702	0.611	0.543
Observations	3696	3696	3696

***, **, and * indicate significance at the 1%, 5%, and 10% level. Z values are reported in parentheses.

measuring the standard deviation of the surface slope. Meanwhile, we take one period lagged smart city construction to reduce the endogenous error. Table 6 represents the endogenous processing result, suggesting that the

economic coordination under the effect of urban surface roughness is more obvious compared with the coefficient of $Lnsprawl_{it}$. In addition, the interaction between urban sprawl and smart city construction is conducive to economic

TABLE 5: Robustness test.

Variable	Impact of smart city	Impact of suburbanization	Impact of interaction
Lnlt_{it-1}	-0.063*** (-7.59)	-0.063*** (-7.59)	-0.063*** (-7.59)
Lnlt_{it-1} (impact)	-0.062*** (-7.56)	-0.065*** (-7.27)	-0.068*** (-5.76)
Lnsmart_{it}	0.001* (1.78)		0.001* (1.88)
$\text{Lnsuburbanization}_{it}$		-0.013*** (-3.62)	-0.011*** (-2.91)
$\text{Lninteraction}_{it}$			0.006* (1.73)
$W * \text{Lnsmart}_{it}$	-0.0007*** (-2.31)		-0.003 (-1.09)
$W * \text{Lnsuburbanization}_{it}$		-0.013 (-1.32)	-0.009*** (-2.04)
$W * \text{Lninteraction}_{it}$			-0.003 (-1.26)
Control variable	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Λ	0.107*** (2.89)	0.113*** (2.87)	0.116*** (2.12)
R^2	0.609	0.612	0.601
Observations	3696	3696	3696

***, **, and * indicate significance at the 1%, 5%, and 10% level. Z values are reported in parentheses.

TABLE 6: Endogenous processing result.

Variable	Impact of smart city	Impact of urban sprawl	Impact of interaction
Lnlt_{it-1}	-0.063*** (-7.59)	-0.063*** (-7.59)	-0.063*** (-7.59)
Lnlt_{it-1} (impact)	-0.060*** (-8.45)	-0.064*** (-7.37)	-0.065*** (-7.16)
Lnsmart_{it-1}	0.001*** (2.36)		0.001* (1.91)
Lnroughness_{it}		0.003* (1.83)	-0.002 (-1.21)
$\text{Lninteraction}_{it}$			0.011* (1.66)
$W * \text{Lnsmart}_{it-1}$	-0.001*** (-2.32)		-0.001 (0.75)
$W * \text{Lnroughness}_{it}$		-0.007*** (-2.28)	-0.004* (-1.66)
$W * \text{Lninteraction}_{it}$			-0.008*** (-2.21)
Control variable	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
λ	0.145*** (4.42)	0.145*** (3.67)	0.142*** (2.67)
R^2	0.609	0.611	0.610
Observations	3696	3696	3696

***, **, and * indicate significance at the 1%, 5%, and 10% level. Z values are reported in parentheses.

growth, as well as coordination, which is consistent with baseline result.

6. Further Research

Some studies have shown that urban sprawl tends to restrict economic growth through changing the level of industrial agglomeration but ignored the effect of construction of smart city (Liangfeng and Li, [31]). According to the studies, urban sprawl weakened the effect of industrial

agglomeration by expanding the scope of space. However, it could be expected that smart city construction could reduce the negative effect of space expansion. We incorporate the geographic change into the construction of agglomeration index, in order to reflect the impact of urban sprawl. We introduce the variables and interactive items to test the interactive effect of industrial agglomeration, urban sprawl, and smart city construction on regional coordination. The benchmark model is revised to the following equation:

$$\text{Lnlt}_{it} - \text{Lnlt}_{it-1} = \lambda W_n (\text{Lnlt}_{it} - \text{Lnlt}_{it-1}) + \gamma \text{Lnlt}_{it-1} + \beta \text{Lnsprawl}_{it} + \kappa \text{Lnagg}_{it} + \varphi \text{Lnsprawl}_{it} * \text{Lnagg}_{it} + \mu_i + \zeta_t + \xi_{it}. \quad (19)$$

The result presented in Table 7 shows that the absolute value of Lnlt_{it-1} has been reduced to a certain extent, when introducing the impact of the dynamic industrial agglomeration. It implies that the regions with higher economies are more attractive to the industrial inflows. In addition, the smart city construction tends to widen the gap among the regions. The reasonable explanation lies in that smart city construction could accelerate the flow of factors and attract

high quality resources in more developed areas. Meanwhile, the interaction term of $\text{Lnsprawl}_{it} * \text{Lnagg}_{it}$ is negative, indicating that the sprawling urban structure weakens the industrial agglomeration effect, which also has attached negative impact on economic growth. The reason lies in that the mismatch between the land supply and factors inflow in the eastern and central western cities has provided reasonable explanation for the result (Lu et al. [32]). The faster

TABLE 7: Mediation effect test from dynamic industrial agglomeration.

Variable	Impact of industrial agglomeration	Interaction effect of smart city construction and agglomeration	Interaction effect of urban sprawl and agglomeration	Interaction effect of triple factors
Lnlt_{it-1}	-0.063*** (-7.59)	-0.063*** (-7.59)	-0.063*** (-7.59)	-0.063*** (-7.59)
Lnlt_{it-1} (impact)	-0.062*** (-6.09)	-0.061*** (-4.92)	-0.062*** (-4.06)	-0.064*** (-3.98)
Lnagg_{it}	0.002** (2.37)	0.001 (1.50)	0.001 (1.37)	0.001 (1.10)
Lnsmart_{it}		0.001* (1.71)		0.001 (1.22)
Lnsprawl_{it}			-0.004* (-1.91)	-0.002 (-1.54)
$\text{Lninteraction}_{it}$		0.0006* (1.82)	-0.007 (-1.06)	0.0003* (1.75)
$W * \text{Lnagg}_{it}$	-0.002 (-1.28)	-0.001 (-1.58)	-0.001 (-1.51)	-0.001* (-1.83)
$W * \text{Lnsmart}_{it}$		-0.001** (-1.98)		0.001* (1.66)
$W * \text{Lnsprawl}_{it}$			0.001 (1.09)	0.001* (1.66)
$W *$		-0.0004 (-0.87)	0.0001** (2.21)	0.0001* (1.77)
$\text{Lninteraction}_{it}$				
Control variable	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
λ	0.152** (2.56)	0.135*** (2.67)	0.102** (2.03)	0.187** (2.34)
R^2	0.521	0.598	0.534	0.492
Observations	3696	3696	3696	3696

***, **, and * indicate significance at the 1%, 5%, and 10% level. T values are reported in parentheses.

TABLE 8: Depreciation rates in different provinces.

Region	Depreciation rate	Region	Depreciation rate (%)
Beijing	9.90	Henan	10.85
Tianjin	10.10	Hubei	10.38
Hebei	10.66	Hunan	9.58
Shanxi	10.28	Guangdong	10.08
Inner Mongolia	10.37	Guangxi	10.49
Liaoning	10.46	Hainan	9.36
Jilin	11.49	Chongqing	9.19
Heilongjiang	10.46	Sichuan	9.53
Shanghai	10.06	Guizhou	9.57
Jiangsu	11.49	Yunnan	9.29
Zhejiang	10.33	Shanxi	9.50
Anhui	10.38	Gansu	9.95
Fujian	10.14	Qinghai	9.46
Jiangxi	10.78	Ningxia	10.36
Shandong	10.93	Average	10.19

TABLE 9: Result of stochastic frontier model.

Variable	Coefficient	Regression	Variable	Coefficient	Regression
$\text{ln}L_{it}$	β_1	0.132*** (0.011)	$\text{tln}L_{it}$	β_8	0.003*** (0.000)
$\text{ln}K_{it}$	β_2	-0.455*** (0.016)	$\text{tln}K_{it}$	β_9	-0.017*** (0.001)
T	β_3	0.018*** (0.003)	σ^2		0.001*** (0.000)
$(\text{ln}K_{it})^2$	β_4	0.385*** (0.062)	η		0.025*** (0.003)
$(\text{ln}L_{it})^2$	β_5	0.003*** (0.001)	γ		0.698*** (0.018)
t^2	β_6	0.001*** (0.000)	Log likelihood		8.979
$\text{ln}K_{it}\text{ln}L_{it}$	β_7	-0.079*** (0.011)	Log likelihood-ratio test of $\text{sigma}-u=0$		7908

***, **, and * indicate significance at the 1%, 5%, and 10% level. Standard deviation values are reported in parentheses.

land supply in the western region weakened the agglomeration effect, which further increased the regional disparity. The interaction of triple factors shows positive effect on economic growth, as well as regional coordination. The reason lies in that smart city construction tends to accelerate the flow of factors, while urban sprawl could weaken the boundary effect created by the distance among the cities.

7. Conclusion and Policy Implications

The rapid improvement in urbanization has promoted discontinuous and low-density spatial expansion. Smart city construction has become an important driving force for the development of urbanization in China. This paper takes the objective and corrected urban night light data as proxy variable of economic growth, combined with LandScan global population dynamics statistics to measure the level of urban sprawl. It is found that urban sprawl is conducive to promoting the coordinated development, which is more obvious among neighboring cities. Further research shows that smart city construction is conducive to economic growth other than regional coordination. The endogenous treatment using urban surface roughness as instrumental variable shows no significant difference. Further research shows that smart city construction together with urban sprawl and dynamic industrial agglomeration could promote economic growth, as well as regional coordination.

The findings are relevant from public policy. The empirical result shows that urban sprawl could strengthen coordination among neighboring cities, but vicious competition relationship restricts the process. Meanwhile, smart city construction is beneficial to economic growth, but not for other cities. There is amount of evidence showing that the formation of the smart infrastructure network has gradually blurred the geographical boundaries among the cities but highlighted the existence of administrative barriers (Tang [33]). Therefore, it is necessary to promote the infrastructure construction, and to explore the feasibility of integrated administration on the basis of integrated infrastructure. The government should strengthen the construction of cooperation mechanism in the process of integration in physical

space and promote the coordinated development among the regions.

The policy implications also lie in improving the market mechanism and continue to deepen the regional coordinated strategy. The empirical results show that there is disconnection between urban sprawl and industrial agglomeration, which is not conducive to economic growth and regional coordination. It suggests that the government should deepen the regional coordinated development strategy, improve the market mechanism, and optimize the land supply structure, so as to promote the coordinated relationship of urban sprawl and industrial agglomeration, in addition to accelerating the cross-regional flow of factors through the construction of smart cities.

Appendix

A. Depreciation rate

Limited by the availability of data, it is difficult to obtain the depreciation rates at the city level. However, taking the same depreciation rate for all cities tends to cause additional error. Therefore, we estimate depreciation rates at the provincial level and take the same value for cities in the same province. According to the “Depreciation Period Table of Fixed Assets of State-owned Enterprises,” we set the depreciation periods of the construction industry and production equipment as forty years and sixteen years, respectively. It is conventionally assumed that the residual value rate of fixed assets is 5%, and the geometric decline calculation formula is specifically expressed as $w_\tau = (1 - \delta)^\tau$, in which $\tau = 0, 1, 2, \dots, T$. The total depreciation rate of each province could be calculated by multiplying the depreciation rate by the corresponding investment weight of construction industry and production equipment. Table 8 shows the result of depreciation rates in different provinces.

B. Technological Progress Rate

We adopt the stochastic frontier model to measure technological progress rate, which could be expressed as transcendental logarithmic production function:

$$\ln Y_{it} = \beta_0 + \beta_1 \ln L_{it} + \beta_2 \ln K_{it} + \beta_3 t + 0.5\beta_4 (\ln K_{it})^2 + 0.5\beta_5 (\ln L_{it})^2 + 0.5\beta_6 t^2 + \beta_7 \ln K_{it} \ln L_{it} + \beta_8 t \ln L_{it} + \beta_9 t \ln K_{it} + v_{it} - u_{it}. \quad (\text{B.1})$$

in which Y is the cities' total output, L indicates the labor force, and K is the cities' capital stock. Both Y and L could be collected in the Statistical Yearbook of Chinese cities, while the capital stock of cities needs to be measured. Perpetual inventory method is used for calculation of capital stock.

(Table 9) shows the regression results of stochastic frontier model, and the technological progress rate could be calculated by equation (B.2).

The total factor productivity, as well as the decomposition indices of scale efficiency, rate of technological

progress, and technological efficiency could be calculated, using Kumbhakar and Lovell's decomposition method, which is expressed as follows:

$$FTP_{it} = \beta_3 + \beta_6 t + \beta_8 \ln L_{it} + \beta_9 \ln K_{it}. \quad (\text{B.2})$$

Data Availability

The underlying data could be found at <https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>. In addition, China industrial enterprise database and China City Statistical Yearbook are used for empirical test. The data could be required through e-mail address: haousts@usts.edu.cn.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Major Project of National Philosophy and Social Science Foundation of China under Grant No. 18ZDA038 and the Project of National Philosophy and Social Science Foundation of Jiangsu Province, China, under Grant No. 20GLC014.

References

- [1] X. Liu, "Spatial efficiency and regional equilibrium: a test of agglomeration effect at the provincial level in China," *The Journal of World Economy*, vol. 37, no. 1, pp. 55–80, 2014.
- [2] X. Liu, S. Li, and M. Qin, "Development lag, market uncertainty and urban sprawl," *Economic Research*, vol. 51, no. 8, pp. 159–171, 2016.
- [3] M. Jiang and Q. Xi, "The measurement of China's major urbanized areas: based on the perspective of population gathering," *Social Sciences in China*, vol. 8, no. 8, pp. 26–46, 2015.
- [4] S. Lv, H. Xue, and K. Wang, "The review and outlook on smart city construction," *Contemporary Economic Management*, vol. 39, no. 4, pp. 53–57, 2017.
- [5] C. Lu and K. Yang, "Construction of evaluation index system of smart city," *Statistics and Decision*, vol. 6, pp. 3–38, 2018.
- [6] Q. Xu, Z. Wu, and L. Chen, "The vision and architecture of a smart city," *Journal of Management Engineering*, vol. 26, no. 4, pp. 1–7, 2012.
- [7] M. Chen, Q. Wang, X. Zhang, and X. Zhang, "Research on the evaluation index system of smart city," *Urban Development Research*, vol. 18, no. 5, pp. 84–89, 2011.
- [8] H. Chourabi, T. Nam, S. Walker et al., "Understanding smart cities: an integrative framework," in *Proceedings of the 2012 45th Hawaii International Conference on System Sciences*, January 2012.
- [9] S. Alawadhi, A. Aldama-NaldaH. Chourabi et al., Building understanding of smart city initiatives," in *Proceedings of the International Conference on Electronic Government*, September 2012.
- [10] E. L. Glaeser and M. E. Kahn, "Chapter 56 sprawl and urban growth," *Handbook of Regional and Urban Economics*, vol. 4, pp. 2481–2527, 2004.
- [11] B. Lee and P. Gordon, "Urban structure its role in urban growth net new business formation and industrial churn," *Region et Development*, vol. 33, pp. 137–159, 2011.
- [12] B. N. Fallah, M. D. Partridge, and M. R. Olfert, "Urban sprawl and productivity: evidence from US metropolitan areas," *Papers in Regional Science*, vol. 90, no. 3, pp. 451–472, 2011.
- [13] S. Farber and X. Li, "Urban sprawl and social interaction potential: an empirical analysis of large metropolitan regions in the United States," *Journal of Transport Geography*, vol. 31, pp. 267–277, 2013.
- [14] G. Di Liddo, "Urban sprawl and regional growth empirical evidence from Italian regions," *Economics Bulletin*, vol. 35, no. 4, pp. 2141–2160, 2015.
- [15] M. Qin and X. Liu, "Does urban sprawl lead to urban productivity losses in China? empirical study based on nighttime light data," *Journal of Finance and Economics*, vol. 41, no. 7, pp. 28–40, 2015.
- [16] S. Sun and G. Wan, "The impact of urban sprawl on the welfare of residents," *Economic Dynamics*, vol. 11, pp. 32–45, 2017.
- [17] X. Jiang, "The impact of urban sprawl and fiscal decentralization on sulphur dioxide emissions-based on the empirical analysis of the panel data of 21 prefecture-level city of Guangdong province from 2007 to 2017," *Public Administration Review*, vol. 12, no. 5, pp. 154–175, 2019.
- [18] Y. Lin and M. Lui, "Growth convergence and income distribution in China," *World Economy*, vol. 8, pp. 3–14, 2003.
- [19] Z. Xu and S. Li, "Analysis on the trend of regional income disparity in China," *Economic Research*, vol. 7, pp. 35–41, 2006.
- [20] C. Tan, Y. Liu, and L. Chao, "Spatial spillovers and the convergence of regional economic growth: a case study based on the Yangtze river delta," *Social Science in China*, vol. 34, no. 3, pp. 76–94, 2012.
- [21] T. He and X. Chen, "Empirical analysis of dynamic externalities and growth convergence among urban agglomeration," *Systems Engineering Theory & Practice*, vol. 37, no. 11, pp. 2791–2801, 2017.
- [22] S. Fu and B. Ren, "Strategic competition, spatial effects and the convergence of China's economic growth," *Economic Dynamics*, vol. 2, pp. 47–62, 2019.
- [23] Y. Yu, "A study on the dynamic spatial convergence of China's provincial total factor productivity," *The Journal of World Economy*, vol. 38, no. 10, pp. 30–55, 2015.
- [24] N. G. Mankiw, D. Romer, and D. N. Weil, "A contribution to the empirics of economic growth," *Quarterly Journal of Economics*, vol. 107, no. 2, pp. 407–437, 1990.
- [25] B. E. López, E. Vayá, and M. Artís, "Regional externalities and growth: evidence from European regions," *Journal of Regional Science*, vol. 44, no. 1, pp. 43–73, 2004.
- [26] J. V. Henderson, T. L. Squires, and A. Storeygard, "The global spatial distribution of economic activity: nature, history and the role of trade," *National Bureau of Economic Research*, vol. 133, no. 1, pp. 357–406, 2016.
- [27] B. Fallah, M. Partridge, and M. R. Olfert, "Uncertain economic growth and sprawl: evidence from a stochastic growth approach," *The Annals of Regional Science*, vol. 49, no. 3, pp. 589–617, 2012.
- [28] C. Xie and M. Wang, "A study on the impact of transportation infrastructure on the spatial distribution of industrial activities," *Management World*, vol. 12, pp. 52–63, 2020.
- [29] H. Yuan and C. Zhu, "Do smart cities accelerate urban innovation?" *China Soft Science*, vol. 12, pp. 75–83, 2020.

- [30] M. Burchfield, H. G. Overman, D. Puga, and M. Turner, "Causes of sprawl: a portrait from space," *Quarterly Journal of Economics*, vol. 121, no. 3, pp. 587–633, 2006.
- [31] H. A. O. Liangfeng and X. Li, "Impact of urban sprawl on regional productivity improvement: based on the perspective of dynamic industrial agglomeration," *Journal of Central South University*, vol. 26, no. 2, pp. 21–31, 2020.
- [32] M. Lu, H. Zhang, and W. Liang, "How does the supply of land in the Midwest boost the wages in the East?" *Social Sciences in China*, vol. 59–83, no. 5, pp. 204–205, 2015.
- [33] W. Tang, "Decentralization, externalities and boundary effects," *Economic Research*, vol. 54, no. 3, pp. 103–118, 2019.

Research Article

Hybrid Algorithm Based on Content and Collaborative Filtering in Recommendation System Optimization and Simulation

Lianhuan Li ¹, Zheng Zhang ², and Shaoda Zhang ³

¹School of International Education, Nanyang Medical College, Nanyang, Henan 473000, China

²School of Computer and Software, Nanyang Institute of Technology, Nanyang, Henan 473000, China

³Cofee Medical Technology Company Limited, Shenzhen, Guangdong 518101, China

Correspondence should be addressed to Zheng Zhang; zhangzheng@nyist.edu.cn and Shaoda Zhang; zhangsd@stu.xmu.edu.cn

Received 12 April 2021; Revised 25 April 2021; Accepted 10 May 2021; Published 19 May 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Lianhuan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper explores and studies recommendation technologies based on content filtering and user collaborative filtering and proposes a hybrid recommendation algorithm based on content and user collaborative filtering. This method not only makes use of the advantages of content filtering but also can carry out similarity matching filtering for all items, especially when the items are not evaluated by any user, which can be filtered out and recommended to users, thus avoiding the problem of early level. At the same time, this method also takes advantage of the advantages of collaborative filtering. When the number of users and evaluation levels are large, the user rating data matrix of collaborative filtering prediction will become relatively dense, which can reduce the sparsity of the matrix and make collaborative filtering more accurate. In this way, the system performance will be greatly improved through the integration of the two. On the basis of the improved collaborative filtering algorithm, a hybrid algorithm based on content and improved collaborative filtering was proposed. By combining user rating with item features, a user feature rating matrix was established to replace the traditional user-item rating matrix. *K*-means clustering was performed on the user set and recommendations were made. The improved algorithm can solve the problem of data sparsity of traditional collaborative filtering algorithm. At the same time, for new projects, it can also predict users who may be interested in new projects according to the matching of project characteristics and user characteristics scoring matrix and generate push list, which effectively solve the problem of new projects in “cold start.” The experimental results show that the improved algorithm in this paper plays a significant role in solving the speed bottleneck problems of data sparsity, cold start, and online recommendation and can ensure a better recommendation quality.

1. Introduction

The recommendation system excavates the resources that users may be interested in or need according to the interest characteristics of different users from the mass information and makes recommendations [1]. As a BI [2] platform based on massive data mining, it is considered to be one of the most effective tools to solve the information explosion. In essence, the recommendation system evaluates the likes of some products that the user has never touched by analyzing the resources selected by the user and feeds back the products with the highest likes among the predicted results to the user.

Nowadays, most recommendation systems tend to sacrifice recommendation quality in order to meet users'

requirements for real-time recommendation. However, the long-term unsatisfactory recommendation quality will reduce users' satisfaction with the website. In addition, some core algorithms of the recommendation system also have problems such as sparse data, narrow application scope [3], and cold start [4]. For a website with a large number of items, users will only have comments on a small part of them, and different users will choose different categories of items for evaluation, which leads to the sparsity of the user evaluation item matrix and seriously affects the quality of recommendation. In addition, the recommendation algorithm is mostly based on the user's evaluation items to make predictions, so for newly registered users, because there is no historical information, the system could not know their

interests and hobbies to make recommendations. For newly launched projects, because no users have made comments on them, they will not be pushed by the system. This is what we call the cold start problem.

In order to solve the above problems and better serve users, this paper studies the recommendation system based on the hybrid algorithm of content and collaborative filtering and its core algorithm, which solves the speed bottleneck and data sparsity problems to a certain extent, ensures the recommendation quality, and promotes the development and application of the recommendation system.

2. Related Work

At present, there are quite a few e-commerce systems using recommendation technology to improve the revenue of enterprises. According to the recommendation technology adopted, the e-commerce recommendation system can be divided into two types: rule-based recommendation system and information filtering recommendation system, and the information filtering recommendation system can be divided into content-based filtering recommendation system and collaborative filtering recommendation system.

The rule-based recommendation systems, such as IBM Web Sphere [5], allow system administrators to formulate rules according to users' static characteristics and dynamic attributes. In essence, a rule is an if-then statement, which determines how to provide different recommendation services under different circumstances. The advantage of a rule-based system is that it is simple and direct, but the disadvantage is that the quality of rules is difficult to guarantee and cannot be updated dynamically. In addition, as the number of rules increases, the system will become more and more difficult to manage. Content-based filtering systems such as Personal WebWatcher [6], Syskill & Webert [7], Letizia [8], Citese [9], IFWEB, WebACE [10], Elfish, and WebPersonalizer use similarity between resources and user interests to filter information. The advantages of the system based on content filtering are simple and effective, but the disadvantages are that it is difficult to distinguish the quality and style of the content of resources, and it could not find new resources that users are interested in, but can only find resources similar to users' own interests. Collaborative filtering systems, such as WebWatcher [11] and Let us Browse [12], use the similarity between users to filter information. The advantage of the collaborative filtering system is that it can find new information of interest for users. Failures are two very difficult problems to solve. One is sparse, that is, when the system starts to use, due to system evaluation, resources have not yet received enough system, and it is difficult to use evaluation to find similar users. The other is scalability. As the number of users and resources of the system increases, the performance of the system will become lower and lower.

The specification of the corpus provided by TREC can be used [13]. They extract feature vectors from the topic description and cases in the document as the initial user demand model and interact with the training set for each topic

to set the initial threshold. Then, you use the test set to determine whether the documents in the test set are greater than a certain topic of the Fujian value condition (if the condition is met). Then, feature vectors are extracted from the positive and negative example document sets to update the initial model, and a filtering system based on the vector model is established [14]. The research on Chinese text filtering technology uses vector space model, demand, and text matching technology based on user demand example keywords in the Chinese text filtering model they proposed and uses vector angle cosine as the similarity coefficient [15]. Based on the text representation of the text, starting from the information of the text logic level, the text structure analysis technology is introduced to improve the matching efficiency of documents and queries in the retrieval of text fragments [16]. In addition, they combined the content-based filtering method with the collaborative filtering method to build a hybrid model of text filtering, so as to make better use of user evaluation information. It proposes an information filtering model based on Bayesian network BMIF, describes the basic structure of information filtering, and provides 6 kinds of nodes to describe the relationship between information filtering events [17]. On this basis, it provides a variety of BMIF usages, uses the vocabulary knowledge represented by BMIF, combines automatic learning with manual interaction, and combines collaborative filtering with content filtering.

The recommendation algorithm, as the key of the recommendation system, is the target of researchers' in-depth study. The excellent algorithm not only requires stable and accurate operation but also must contact the application environment with a certain universality, so it can meet the needs of the user key to see the effective use of the algorithm, because the method recommended by the algorithm will be different, which requires researchers to study rational selection through the experiment and targeted trade-off for the field. After 20 years of development of the recommendation system, scholars have made use of knowledge in different fields to improve the recommendation algorithm from multiple perspectives and put forward different recommendation algorithms [18]. At present, the recognized recommendation algorithms include collaborative filtering recommendation, content-based recommendation, knowledge-based recommendation, and hybrid recommendation algorithm [19]. The collaborative filtering recommendation algorithm is the earliest recommendation algorithm [20]. The characteristic of this algorithm is to find similar users of target users and then recommend products or information for target users by using similar users. The recommendation advantage of the collaborative filtering algorithm is that, under the condition of relatively dense rating data, collaborative filtering can bring better recommendation effect for users than other algorithms. However, the resource required by this algorithm is the user-project rating data, so it does not need to know the user information or make clear the project information. Because the information of the user and the project is not easy to obtain, the algorithm has a low complexity table and a wide range of application fields [21]. However, the reliance on scoring also becomes a defect of the

collaborative filtering algorithm. When the scoring matrix is sparse, the recommendation effect is greatly reduced. In view of the imperfection of the collaborative filtering recommendation algorithm, scholars began to improve collaborative filtering from multiple perspectives. The LDA topic model was introduced into the collaborative filtering algorithm [22], and the potential factor vectors of users and projects were obtained by using the LDA topic model. The sparse matrix was reduced by dimensionality reduction of the scoring matrix, and then, the similarity between users and projects was calculated in the low-dimensional space. Collaborative filtering is improved by defining the similarity of user attributes in social networks [23] and giving the calculation method of attribute similarity. Considering the length and dimension of the space vector comprehensively, the improved cosine similarity is used to find similar users [24]. Based on practical experience, the expert trust factor of users is introduced into the algorithm [25]. Considering the structural relationship between users, a modeling method for time-series behavior among users is proposed [26]. Some scholars seek similar users by effectively analyzing the trust relationship between users [27]. Some other scholars use relevant theories and methods to cluster items and predict unrated items in the same category to densify the scoring matrix in order to mitigate the adverse impact of scoring sparsity on the collaborative filtering algorithm [28].

Aiming at the problems of the current recommendation system and combining the advantages and disadvantages of content filtering and collaborative filtering, a content-based filtering is proposed. In the hybrid mode based on the combination of collaborative filtering recommendation technology, the recommendation filtering technology is fully utilized. The advantages of the two recommendation technologies make up for each other's shortcomings, improve the performance of the recommendation system, help companies improve the quality of customer service, enhance the cross-selling capabilities of the e-commerce system, improve corporate competitiveness, and provide customers with more accurate and personalized services. Using the collaborative filtering recommendation method recommended by the user data set and then based on the content reuse recommendation method, the recommended data set is filtered by establishing a model interested in the recommended data set. In order to realize that the content-based recommendation method is used to optimize the results of collaborative filtering recommendation, the purpose of this hybrid is recommended, but the purpose is very strong.

3. Hybrid Recommendation Algorithm and Recommendation Strategy Based on Content and Collaborative Filtering

3.1. Research on the Hybrid Recommendation Model Based on Content and Collaborative Filtering. The schema framework of the hybrid mode recommendation algorithm based on content and collaborative filtering is shown in Figure 1.

The whole recommendation is divided into two modules, namely, content filtering recommendation

module and collaborative filtering recommendation module, both of which are invisible to users. The preparation process of the dataset of the recommendation algorithm is as follows: first, the user's interest is extracted from the shopping history data of the user and the topic vector and feature vector preprocessed by the network log, and the data processing is established based on the recommendation module of content filtering. Then, based on user interest characteristics, user rating data, and current access sequence data, a recommendation module based on collaborative filtering is constructed to extract the nearest neighbors of the user and the nearest neighbors of the current access sequence (item). Then, it integrates two recommendation weighted sum calculation modules for the similarity calculation model of mixed recommendation (i.e., recommendation processing) to generate the recommended top visit sequence. The web server recommends the sequence to the user and accesses the recommendation sequence on the user, the adaptive adjustment of the recommended model, and the idle speed value of the feedback information to obtain the best recommendation data.

In order to realize personalized recommendation service, we must first collect the user's personal information and establish the user's interest characteristic description model. The ratio of the time spent browsing a web page to the number of characters on the page can effectively reveal the user's interest, which is related to the categories of information, and these categories are determinable and relatively stable. Users browse the Web information including each page clicks, stay time, access, and order, and each page URL can be found in the Log of the proxy server, and the user visited Web pages can be found in the Cache of the proxy server, so you can go through the Web mining way to get the user's interest.

The optimal feature items are those words with the largest mutual information amount with the related text set $Rel(Q)$. The logarithmic mutual information amount between the words and the related text set is calculated by the following formula:

$$\log NI(\alpha, rel(Q)) = \log \left(Q \left(\alpha | \alpha \in \frac{rel(Q)}{Q(\alpha)} \right) \right). \quad (1)$$

The cosine similarity between user preference document and project document is

$$\text{sim}(\beta, e_j) = \cos(e_j, e_c) = \frac{\sum_i \alpha_i \varepsilon_{ij}}{\sqrt{\alpha_i^2} \sqrt{\varepsilon_{ij}^2}} \quad (2)$$

The higher the calculated similarity, the more preference the users have for this feature. The biggest problem facing TF-IDF is the choice of features. The content category of users is based on the similarity between user interests, that is, the similarity between user feature vectors. Here, the commonly used method of cosine of included angle is selected. The similarity of user interests is

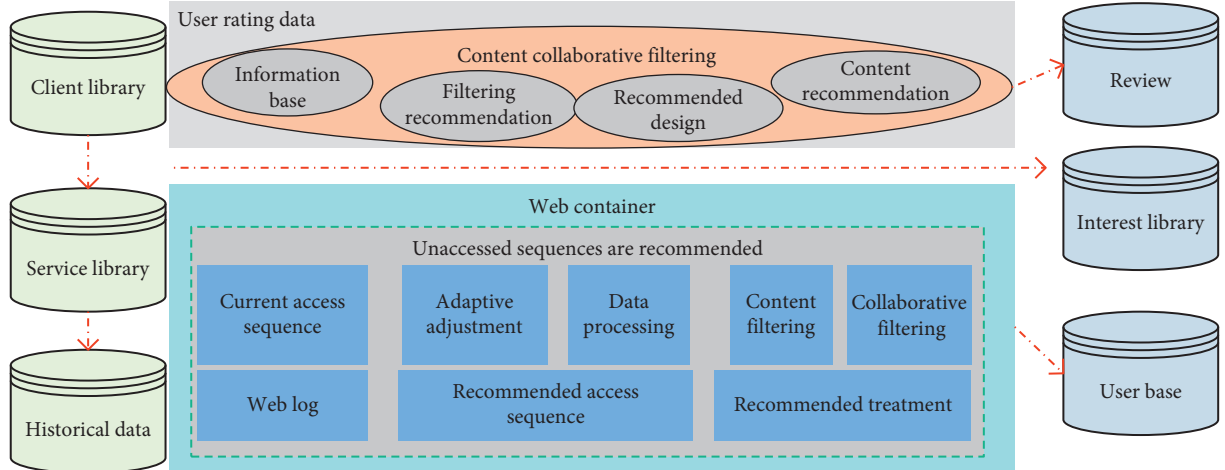


FIGURE 1: Schematic diagram of a hybrid recommendation model based on content and collaborative filtering.

$$\text{sim}(U, V) = \sum_i (u_i, v_i) * \frac{1}{\sqrt{u_j^2} \sqrt{v_i^2}} \quad (3)$$

Clustering is carried out according to the similarity between user feature vectors so that users with similar interests can be classified into one group for easy processing. Meanwhile, for new product information documents, a list of recommended users can be obtained by judging their categories. It is assumed that the classification of user sets is controlled manually, so the recommender system clustering method can be adopted.

After the initial recommendation model has been established and the threshold has been set, the similarity between the text about the product introduction in the product information database and the model vector of a user's interest topic can be calculated. If the similarity value is greater than or equal to fujian, it is considered to be associated with the user's interests. You can recommend to

the user through the Web server and then determine whether the recommendation is valid. The user can adaptively modify the model vector or adjust the proximity value according to the judgment result so that the recommended performance can be realized. The system is continuously improved to better serve the user.

If the recommended access sequence is judged to be related to the user's interest, it will browse its relevant information, and then, the recommended access sequence will become the current access sequence. When adjusting the model vector, the interest topic vector can be extracted from the current access sequence, and the feature vector can be extracted from the user's shopping history data and the Web log (the Web log has changed accordingly). The new model vector is obtained by the weighted sum operation of the topic vector and feature vector. Let the weights be A' , B' , C' , and D' , respectively, then

$$pf(Q) = A' * pf_0(Q) + B' * pf_1(Q) + C' * pf_2(Q) + D' * pf_3(Q). \quad (4)$$

Further, the more the items are rated by two users, the more similar their preferences would be if both users rated them higher or lower than the average rating of the two users. To put it simply, the more items both users like and dislike, the more similar the two users' interests and preferences are to some extent. For example, Table 1 shows the rating of some users on items. The scoring rule is 5-point scale, and the higher the rating, the greater their preference for the item. Now, let us calculate the rating prediction of User3 on Item3.

We consider a user's rating of an item greater than or equal to its average rating as a favorite and a user's rating less than its average rating as a dislike. The reason why the average score is chosen is that some users' score above 3 points means they like it, while some users' score above 4 points means they like it. Therefore, compared with the average score, it can better reflect the user's liking degree. Therefore, user interest similarity can be integrated into the

traditional similarity calculation, and the improved similarity formula of user u and user v is defined as follows:

$$\text{sim}(u, v) = \alpha * \text{PCsim}(u, v) + (1 - \alpha) * \text{Interest}(u, v). \quad (5)$$

3.2. Improved Content and Collaborative Filtering Algorithm Recommendation System Based on K-Means Clustering. A collaborative filtering recommendation algorithm based on K-means clustering is proposed. The new algorithm has two components: offline and online. When offline, first of all, users are clustered according to their characteristic data to form several clustering clusters. When online, the clustering cluster to which the target user belongs to is determined according to the similarity between the target user and each clustering center, so as to find the nearest neighbor in the cluster. Then, based on the preference of the nearest

TABLE 1: User rating data.

	Item 1	Item2	Item3	Item4	Item5
User1	3	4	4	6	
User2	2	5	2	4	2
User3	1	3	1	3	2

neighbor group to the project, we can predict the interest preference of the target users and finally get the recommendation.

The specific idea is to apply K -means clustering to collaborative filtering. For the whole user space, the similarity between users and the clustering center is calculated according to users' purchasing habits and scoring characteristics (that is, the user-item scoring matrix), and the clustering cluster is assigned to users according to the principle of nearest distance, thus the whole user space can be divided into several small groups. Based on the scoring characteristics of all users in each cluster, a virtual user is generated for each cluster. As the representative of all users in the cluster, the rating of the virtual user to the project can be the average of all users in the cluster to the project. At this point, the project ratings of all virtual users form a new search space (virtual user-project rating matrix), which replaces the original user-project rating matrix. When online recommendation is made, it only needs to calculate the similarity between target users and all virtual users, determine the cluster to which the target users belong according to the similarity level, search for neighbors in the cluster, and generate recommendation. The algorithm flow is shown in Figure 2.

For a large recommendation system, there will be a lot of user and project data. However, only a small fraction of the total project space has been evaluated by users, which is known as the data sparsity problem. By clustering, the data dimension can be greatly reduced. After neighborhood users are identified, the degree of preference of target users for unrated items can be predicted based on neighborhood users' preferences. The prediction scoring formula for the project is as follows:

$$\text{pre} = \gamma_{\alpha} + \frac{\sum (\gamma_{ij} - \gamma_i) * \text{sim}(\alpha, i)}{\sum \text{sim}(\alpha, i)} \quad (6)$$

When the hybrid recommendation algorithm starts to run, this article first uses the judgment conditions to process the user score data on the item. The total score of the project is less than 20 users. It is recommended to use them directly based on the content filtering method, which is also considered to be collaborative. The filtering recommendation algorithm for low-scoring data is effective in the fact that there are few user recommendations. The content-based recommendation algorithm is directly used to recommend similar items for users through item feature matching. This will result in mediocre recommendations, but it will also avoid the risk of invalid recommendations due to collaborative filtering of similar users not being able to find them. Of course, the value of 20 here is not fixed. In practical application, it can be adjusted according to the situation.

In addition, in the process of establishing the hybrid algorithm model proposed in this paper, the calculation

method is not rigid with the traditional algorithm, but improved or innovated on the basis of the traditional calculation method, which is mainly reflected in the following points:

- (1) A method to optimize the user similarity calculation formula by using project heat was proposed
- (2) In order to present users' preferences more stereoscopic, the table-oriented feature extraction is carried out in the content-based recommendation algorithm, and the square-one method for calculating users' similarity using the interest model is presented
- (3) According to the characteristics of the algorithm in this paper, a method to derive the weight coefficients of different features by using variance is proposed

The purpose of the content-based recommendation algorithm is to effectively filter out the third category of users whose interests are different from those of the target users, and the work required in this process generally includes three steps. The first step is to extract project features. The second step is to establish the user interest model.

3.2.1. Extract Project Features. For nonliterary items, it is a difficulty to propose item features, but it is also the key to the effectiveness of the algorithm. Tags have the characteristics of complete classification and concise description. Therefore, the choice of labels as the feature description of the project can mark the characteristics of the project well from the point of view of user preferences and can also properly reflect the user's interest demands. In order to better describe the project, this paper first extracts a number of general attributes of the project, then, on this basis, refines the different attributes, and gets the different project characteristics of each attribute; the same kind of project features are included in the same project attribute.

After extracting the features of the project, the next step is to construct the user's interest model, which can also be said to be the interest vector that represents the user's interest and is composed of the project attributes and the project features contained in the attributes.

3.2.2. Establish User Interest Model. The user's interest model is composed of the project attributes and the features contained in the attributes, so for the user, what kind of project can be used to establish their interest model is a question. At this point, the user-project scoring matrix comes in handy. To build a user interest model, it is necessary to count the items that users like and extract their signature features.

3.2.3. Calculate the Similarity between Users. After extracting item features and establishing the interest model for users by using item features and scoring matrix, the next step is to consider how to use the user interest model to calculate the similarity between users. The calculation of similarity can often be transformed into distance visually.

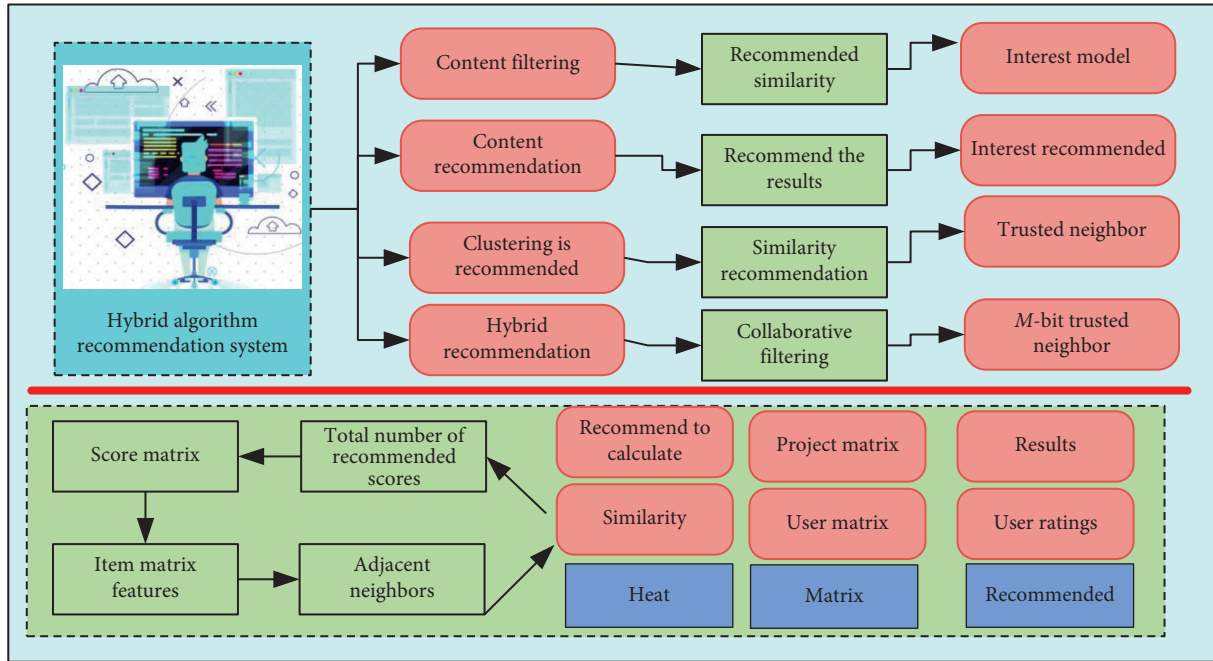


FIGURE 2: Recommendation system optimization diagram of the hybrid algorithm.

Euclidean distance can be used to calculate the distance between points in the multidimensional space.

Replace all the features of the items that users like in the user's interest model with the corresponding number of movies containing this feature. That is, a new representation of the user interest model is obtained by summarizing the user-item attribute preference vector directly. The reason for this is that you do not need to consider what features the user prefers in the following steps, just what the user likes and dislikes about the project. The new form is shown in Table 2.

4. Test Experiment

In this paper, we use the MovieLens dataset and the scientific literature experimental dataset to test the performance of the mixed mode recommendation algorithm, the user-based collaborative filtering recommendation algorithm, and the content filtering recommendation algorithm, respectively.

We selected 12,500 scoring data from the user rating database as the experimental dataset, which contained 248 users and 1120 movies in total, among which each user rated at least 20 movies, with the score value being an integer from 1 to 5. The higher the value, the higher the user's preference for the movie. Each record in the dataset describes information such as user ID, item ID, user rating value of the item, and timestamp. Different user ID numbers represent different users, and different item ID numbers represent different movies.

In order to measure the sparsity of the entire dataset, we introduce the concept of the sparsity level, which is defined as the percentage of unrated items in the user rating data matrix.

Experiment 1. Using the MovieLens dataset, check the performance comparison of the two algorithms under

different training set test set proportions, that is, considering different data sparsity degrees. The number of nearest neighbor users is set to 30. The experimental results are shown in Table 3 and Figure 3.

From the experimental results, the average absolute deviation of the MAE value of the mixed mode recommendation algorithm is smaller than that of the collaborative filtering recommendation algorithm based on user values. In the line chart of the MAE line of the mixed mode recommendation algorithm based on the user collaborative filtering algorithm, it is below the MAE line. As the training set data increases, the gap between the two continues to shrink. This means that the recommendation effect of the mixed mode recommendation algorithm is better than that of the recommendation algorithm based on user collaborative filtering, but this advantage decreases with the increase of the training set.

Experiment 2. Using the MovieLens dataset, check the performance comparison of the two algorithms under the condition that the proportion of the training set test set is constant, i.e., the sparsity degree, and the size of the nearest neighbor set is different, and consider the influence of the size of the nearest neighbor set on the algorithm performance.

When the partition coefficient $x = 0.8$ was selected (it can be seen from Experiment 1 that the prediction effect was better when $x = 0.8$), the experimental results are shown in Table 4 and Figure 4.

It can be found from the experimental results that the average absolute deviation MAE value of the mixed mode recommendation algorithm is smaller than the MAE value of the recommendation algorithm based on user collaborative filtering, and the MAE broken line of the mixed mode

TABLE 2: Numerical representation of the user interest model.

Number of attributes/features/users	Attribute 1	Attribute 2	...	Attribute m
User 1	23 13 9 6	21 13 14 8	...	6 2 3 4
User 2	17 11 12 8	12 13 9 8	...	6 4 2 4
...
User n	14 9 7 2	8 5 5 3	...	6 7 5 4

TABLE 3: Experimental results of different data sparsity degrees.

Proportion of training set (x)	Mean absolute deviation MAE	
	Recommendation algorithm based on user collaborative filtering	Mixed mode recommendation algorithm
0.2	0.8345	0.7868
0.3	0.7981	0.7653
0.4	0.7753	0.7575
0.5	0.7591	0.7382
0.6	0.7493	0.7297
0.7	0.7383	0.7242
0.8	0.7304	0.7204
0.9	0.7252	0.7144

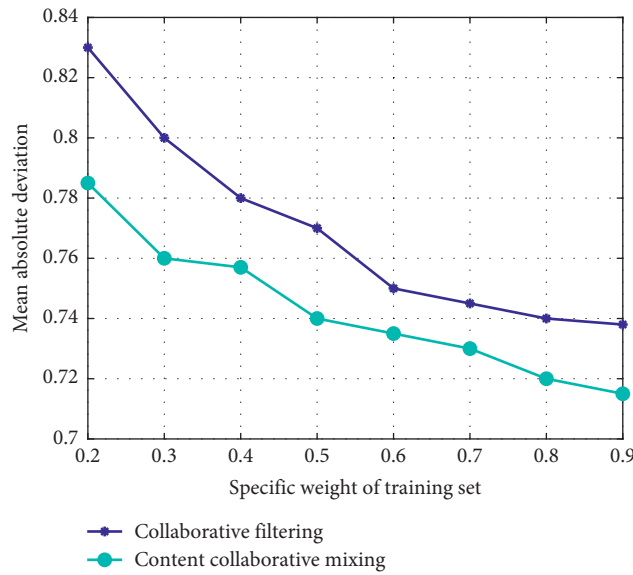


FIGURE 3: Experimental results of different data sparsity degrees.

TABLE 4: Experimental results of different sizes of nearest neighbor sets.

Proportion of training set (x)	Mean absolute deviation MAE	
	Recommendation algorithm based on user collaborative filtering	Mixed mode recommendation algorithm
10	0.7265	0.7163
20	0.7233	0.7117
30	0.7212	0.7103
40	0.7235	0.7115
50	0.7274	0.7152
60	0.7322	0.7227
90	0.7451	0.7284
160	0.7263	0.7164

recommendation algorithm is also below the MAE broken line of the recommendation algorithm based on user collaborative filtering, which also indicates that, under the

condition of the same data sparse degree, the recommendation effect of the hybrid mode recommendation algorithm is better than that of the user collaborative filtering

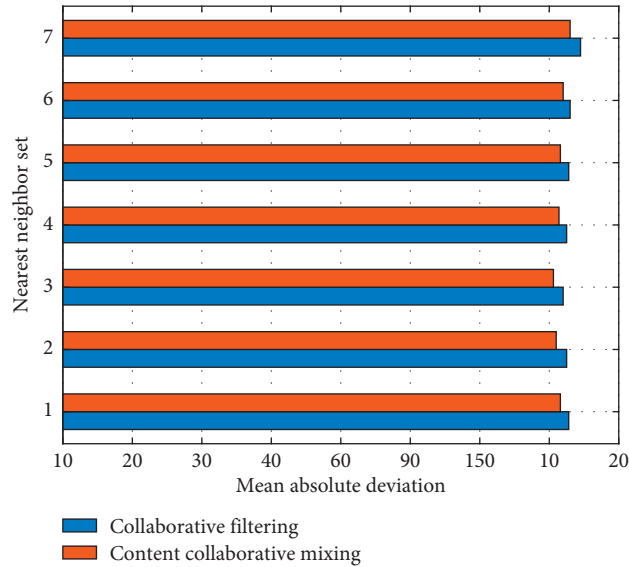


FIGURE 4: Experimental results of different sizes of nearest neighbor sets.

recommendation algorithm. With the increase of the number of the nearest neighbor sets, the recommendation effect will decrease, indicating that the more accurate prediction value can be obtained within a small range of the nearest neighbor values.

In the optimization process of the nearest neighbor selection, the filter parameters of similarity are set to eliminate the users or items that are not very similar, and then, the best value of the filter parameters is determined by combining the improvement of the score prediction part. Figure 5 shows the effect of similarity filter parameters on recommendation quality.

As can be seen from Figure 5, if the filter value of similarity is too small, that is, the filter parameter of similarity is too small and does not filter the nearest neighbor, it does not significantly improve the recommendation performance. With the increase of similarity filtering, the MAE value gradually increases, indicating that the similarity filtering parameter has gradually started to play a role in filtering out users who are not very similar. When the parameter is set to 5, the MAE value reaches the maximum and then begins to decline, that is, the recommendation performance is getting higher and higher. The collaborative filtering recommendation algorithm based on content recommendation quality is slightly higher than that of the collaborative filtering recommendation method. If the content filter-based recommendation method is not used in the case of “false neighbors,” it is also effective to find similar users by optimizing the Pearson correlation coefficient; therefore, introducing heat project to optimize the Pearson correlation coefficient method is effective.

The following four experiments were used to compare the performance of the three algorithms under different sparsity degrees, and the training set and test set were randomly selected from the dataset according to a certain proportion. Here, the number of nearest neighbors is chosen as 4 to test the performance of MAE values of the three algorithms under different sparsity.

Figure 6 shows the comparison of the recommendation quality of the three algorithms under four experiments. Green represents user-based CF, blue represents item-based CF, and red represents the hybrid algorithm proposed in this paper. As can be seen from the figure, under the same other conditions, MAE values of the three algorithms are continuously decreasing with the decrease of dataset sparsity in the four experiments, that is, recommendation quality is continuously increasing. If we carefully look at the figure above, we can find that, in Experiments 1 and 2, the optimization degree of the hybrid algorithm proposed in this paper is higher than that in Experiments 3 and 4. That is to say, the more sparse the dataset is, the more obvious the advantage of using the hybrid algorithm is.

It serves as a reference in the collaborative filtering algorithm and fixes the relationship between M and N . Assuming that compared with the number of neighbors [FI N and the change of parameter M], for $M = 2N/3$, the collaborative filtering recommendation algorithm and the proposed hybrid recommendation method have a change in recommendation effect. In addition to comparing the effects of hybrid algorithms and collaborative filtering algorithms, the purpose of this step is also recommended. It also includes finding the optimal value of the number of users N as a fixed value, so as to obtain the relationship between M and N through further experiments. The experimental results are shown in Figure 7.

As can be seen from Figure 7, the recommendation accuracy of the hybrid algorithm proposed in this paper is better than that of the collaborative filtering algorithm. Moreover, with the increasing number of neighbors, the recommendation quality of the two recommendation algorithms both shows a trend of first increase and then decrease. When $N = 4$, MAE of the collaborative filtering recommendation algorithm is the minimum, and the quality of the recommendation algorithm is the best. When the relationship between M and N is fixed and $M = 2N/3$, the recommended quality is the best. The

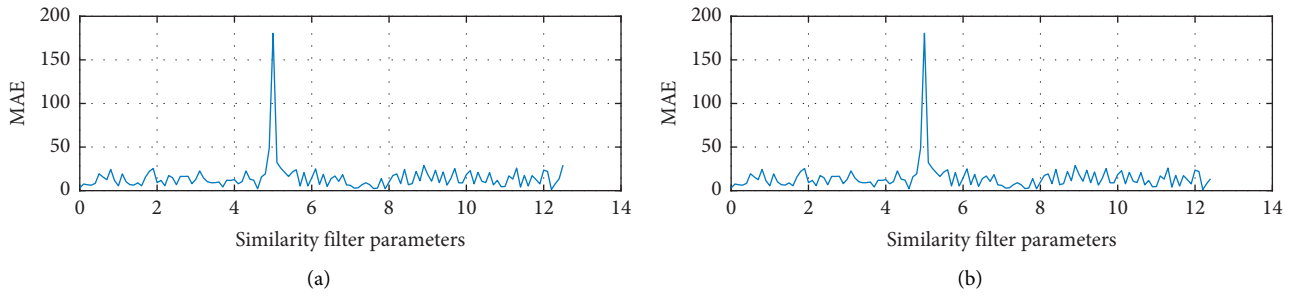


FIGURE 5: Influence of similarity filtering on user recommendation quality.

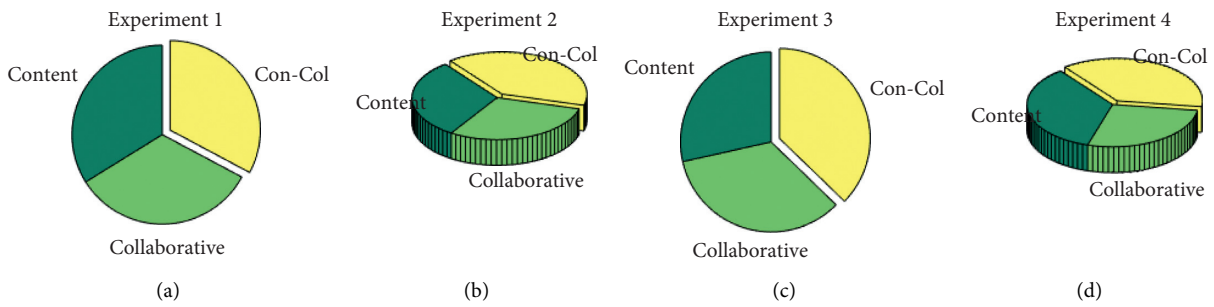


FIGURE 6: Performance of algorithms with different sparsity degrees. (a) Experiment 1. (b) Experiment 2. (c) Experiment 3. (d) Experiment 4.

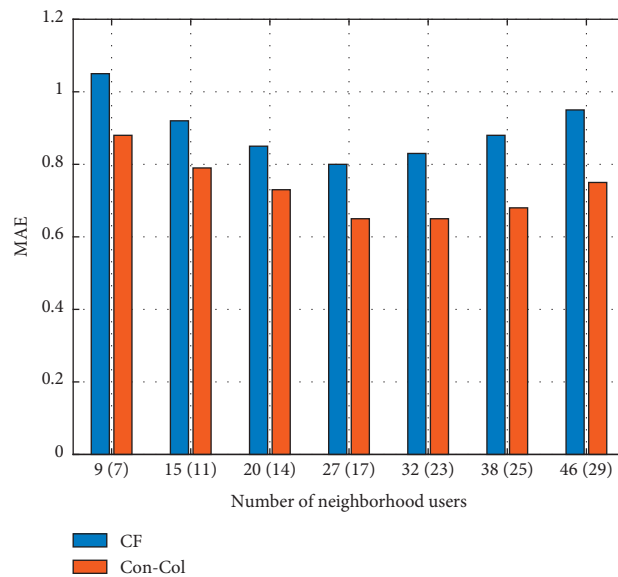


FIGURE 7: The trend of recommended quality when the number of neighbors changes.

reasonable selection of the number of similar users has great influence on the recommendation quality of the recommendation algorithm.

As can be seen from Figure 8, among the three mixed recommendation algorithms, the recommendation algorithm proposed in this paper has the lowest MAE value, that is, the validity of the recommendation algorithm is the best.

Moreover, with the increasing number of neighbors N , the advantage becomes more and more obvious. In addition, after the MAE of the three algorithms reaches the lowest, the IVIAE bar graph of the hybrid algorithm proposed in this paper rises more gently, indicating that compared with the other two recommendation methods, the hybrid recommendation algorithm proposed in this paper is insensitive to

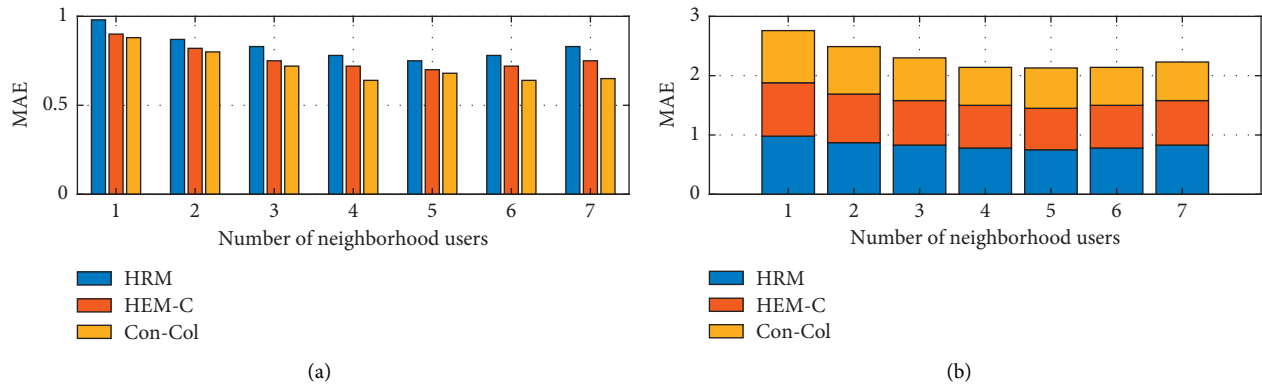


FIGURE 8: Comparison of recommendation validity of mixed recommendation methods.

the change of the number of neighbors, and the recommendation is more stable.

5. Conclusion

Combining the advantages and disadvantages of content filtering and collaborative filtering, we proposed a hybrid recommendation technology based on content filtering and collaborative filtering and studied the workflow, user characteristic description, data processing algorithm, and recommendation strategy of the hybrid recommendation technology in detail. This method not only makes use of the advantages of content filtering but also can carry out similarity matching filtering for all items, especially when the items are not evaluated by any user and can be filtered out and recommended to users, avoiding the problem of early level. At the same time, this method also takes advantage of the advantages of collaborative filtering. When the number of users and evaluation levels are large, the user rating data matrix of collaborative filtering prediction will become more dense, which can reduce the sparsity of the matrix and make collaborative filtering more accurate. The performance of the mixed mode recommendation technology based on content and collaborative filtering is verified. We designed experiments using the MovieLens dataset and the scientific literature experiment dataset to test the system performance of the mixed mode recommendation technology, user-based collaborative filtering recommendation technology, and content-based filtering recommendation technology, respectively. The experimental data show that the performance of the hybrid mode recommendation technology based on content filtering and collaborative filtering is better than that of the two technologies.

Data Availability

No data were used to support this study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by Henan Soft Science Research Program Project (212400410192): Research on the Application of Recommendation Algorithm Based on Multivariate Collaborative Filtering in Medical Practice Qualification Examination.

References

- [1] R. N. Behera and S. Dash, "A particle swarm optimization based hybrid recommendation system," *International Journal of Knowledge Discovery in Bioinformatics*, vol. 6, no. 2, pp. 1–10, 2016.
- [2] T. Horváth and A. de Carvalho, "Evolutionary computing in recommender systems: a review of recent research," *Natural Computing*, vol. 16, pp. 441–462, 2017.
- [3] B. Song, Y. Gao, and X.-M. Li, "Research on collaborative filtering recommendation algorithm based on mahout and user model," *Journal of Physics: Conference Series*, vol. 1437, no. 1, pp. 012095–012101, 2020.
- [4] J. Li and Z. Ye, "Course recommendations in online education based on collaborative filtering recommendation algorithm," *Complexity*, vol. 2020, Article ID 6619249, 10 pages, 2020.
- [5] L. Xu and X. Lu, "Collaborative filtering recommendation system based on interest drifting and similar cloud," *Journal of Liaoning Technical University (Natural Science Edition)*, vol. 36, no. 7, pp. 761–766, 2017.
- [6] G. Geetha, M. Safa, C. Fancy, and D. Saranya, "A hybrid approach using collaborative filtering and content based filtering for recommender system," *Journal of Physics: Conference Series*, vol. 1000, pp. 012101–012123, 2018.
- [7] L. Zuping, "Collaborative filtering recommendation algorithm based on user interests," *International Journal of U- & E-Service*, vol. 8, no. 4, pp. 311–319, 2015.
- [8] J. Jooa, S. Bangb, and G. Parka, "Implementation of a recommendation system using association rules and collaborative filtering," *Procedia Computer Science*, vol. 91, pp. 944–952, 2016.
- [9] R. Chen, Q. Hua, Y.-S. Chang, B. Wang, L. Zhang, and X. Kong, "A survey of collaborative filtering-based recommender systems: from traditional methods to hybrid methods

- based on social networks,” *IEEE Access*, vol. 6, pp. 64301–64320, 2018.
- [10] Z. Qianqian and W. Dongliang, “A SVM recommendation IoT model based on similarity evaluation and collaborative filtering of multi-angle knowledge units,” *International Journal of Computers & Applications*, vol. 42, no. 3, pp. 278–281, 2018.
- [11] Y. Yoo, J. Kim, and B. Sohn, “Evaluation of collaborative filtering methods for developing online music contents recommendation system,” *Transactions of the Korean Institute of Electrical Engineers*, vol. 66, no. 7, pp. 1083–1091, 2017.
- [12] X. Song, Y. Guo, Y. Chang et al., “A hybrid recommendation system for marine science observation data based on content and literature filtering,” *Sensors*, vol. 20, no. 22, pp. 6414–6428, 2020.
- [13] Y. Xu and N. Zhu, “Hybrid recommendation algorithm based on long-term and short-term interest and matrix factorization for collaborative filtering,” *Journal of Physics: Conference Series*, vol. 1624, no. 4, pp. 042015–042021, 2020.
- [14] X. Liu, “A collaborative filtering recommendation algorithm based on the influence sets of e-learning group’s behavior,” *Cluster Computing*, vol. 22, no. 5, pp. 121–131, 2017.
- [15] J. Shi, Y. Tang, and P. Xu, “A collaborative filtering recommendation method based on clustering and user trust,” *Journal of Computational Information Systems*, vol. 11, no. 18, pp. 6845–6852, 2015.
- [16] Y. Gao, C. Huang, M. Hu, J. Feng, and X. Yang, “Research on book personalized recommendation method based on collaborative filtering algorithm,” *IOP Conference Series: Earth and Environmental Science*, vol. 252, no. 5, pp. 052099–052105, 2019.
- [17] H. Xia, J. J. Li, and Y. Liu, “Collaborative filtering recommendation algorithm based on attention GRU and adversarial learning,” *IEEE Access*, vol. 8, pp. 208149–208157, 2020.
- [18] Z. Romadhon, “Various implementation of collaborative filtering-based approach on recommendation systems using similarity,” *Kinetik Game Technology Information System Computer Network Computing Electronics and Control*, vol. 5, no. 3, pp. 134–146, 2020.
- [19] M. Hikmatyar, “Book recommendation system development using user-based collaborative filtering,” *Journal of Physics: Conference Series*, vol. 1477, no. 3, pp. 032024–032032, 2020.
- [20] S. Hattori, R. Misawa, H. Ishikawa, and Y. Takama, “Hybrid recommender system based on collaborative filtering employing personal values-based user model,” *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, vol. 29, no. 4, pp. 628–636, 2017.
- [21] Y. Yan and H. Xie, “Collaborative filtering recommendation algorithm based on user preferences,” *Journal of Physics: Conference Series*, vol. 1549, no. 3, pp. 032147–032155, 2020.
- [22] N. Ifada, N. F. D. Putri, and M. K. Sophan, “Normalization based multi-criteria collaborative filtering approach for recommendation system,” *Rekayasa*, vol. 13, no. 3, pp. 234–239, 2020.
- [23] J. Deng, J. Guo, and Y. Wang, “A novel K-medoids clustering recommendation algorithm based on probability distribution for collaborative filtering,” *Knowledge-Based Systems*, vol. 175, no. 7, pp. 96–106, 2019.
- [24] T. Zhang, “Research on collaborative filtering recommendation algorithm based on social network,” *International Journal of Internet Manufacturing and Services*, vol. 6, no. 4, pp. 343–256, 2019.
- [25] P. K. Singh, P. K. D. Pramanik, and P. Choudhury, “An improved similarity calculation method for collaborative filtering-based recommendation, considering neighbor’s liking and disliking of categorical attributes of items,” *Journal of Information and Optimization Sciences*, vol. 40, no. 2, pp. 397–412, 2019.
- [26] J. Wang, A. K. Sangaiah, and W. Liu, “A hybrid collaborative filtering recommendation algorithm: integrating content information and matrix factorisation,” *International Journal of Grid and Utility Computing*, vol. 11, no. 3, pp. 367–378, 2020.
- [27] B. S. Neysiani, N. Soltani, N. Soltani, R. Mofidi, and M. H. Nadimi-Shahraki, “Improve performance of association rule-based collaborative filtering recommendation systems using genetic algorithm,” *International Journal of Information Technology and Computer Science*, vol. 11, no. 2, pp. 48–55, 2019.
- [28] J. Yamawaki, Y. Kudo, and T. Murai, “Recommendation method based on interrelationship mining and collaborative filtering,” *Transactions of Japan Society of Kansei Engineering*, vol. 17, no. 4, pp. 481–488, 2018.

Research Article

Research on Multi-Target Network Security Assessment with Attack Graph Expert System Model

Yunpeng Li  and Xi Li 

Qingdao Vocational and Technical College of Hotel Management, Qingdao, Shan Dong 266100, China

Correspondence should be addressed to Yunpeng Li; liyunpeng@qchm.edu.cn

Received 31 March 2021; Revised 20 April 2021; Accepted 29 April 2021; Published 7 May 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Yunpeng Li and Xi Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of the Internet, network attacks often occur, and network security is widely concerned. Searching for practical security risk assessment methods is a research hotspot in the field of network security. Network attack graph model is an active detection technology for the attack path. From the perspective of the attacker, it simulated the whole network attack scenario and then presented the dependency among the vulnerabilities in the target network in the way of directed graph. It is an effective tool for analyzing network vulnerability. This paper describes in detail the common methods and tools of network security assessment and analyzes the construction of theoretical model of attack graph, the optimization technology of attack graph, and the research status of qualitative and quantitative analysis technology of attack graph in network security assessment. The attack graph generated in the face of large-scale network is too complex to find the key vulnerability nodes accurately and quickly. Optimizing the attack graph and solving the key attack set can help the security manager better understand the security state of the nodes in the network system, so as to strengthen the security defense ability and guarantee the security of the network system. For all kinds of loop phenomena of directed attribute attack graph, the general method of eliminating loop is given to get an acyclic attack graph. On the basis of acyclic attack graph, an optimization algorithm based on path complexity is proposed, which takes atomic attack distance and atomic weight into consideration, and on the basis of simplified attack graph, minimum-cost security reinforcement is carried out for the network environment. Based on the ant colony algorithm, the adaptive updating principle of changing pheromone and the local searching strategy of the adaptive genetic algorithm are proposed to improve the ant colony algorithm. The experimental results show that compared with the ant colony algorithm, the improved ant colony algorithm can speed up the process of solving the optimal solution. When the number of attack paths is large, the advantages of the improved ant colony algorithm in solving accuracy and late search speed are more obvious, and it is more suitable for large-scale networks.

1. Introduction

With the increasing number of hosts accessing the Internet, coupled with the complexity of various application hardware and software as well as various network protocols, some code loopholes will occur in the stage of design, development, deployment, and maintenance. In a sense, these loopholes will cause the vulnerability of network system. In addition, the network security tools update too slowly, and hackers use the new technology to network environment vulnerabilities to attack. It also makes the network possible to be paralyzed. The outbreak of all kinds of virus events is a wake-up call for

us. With the improvement in the education level of hackers and the convenient network communication and other modern characteristics, there are more and more systematic attacks against valuable information [1–3].

At first, people use network security tools for network defense. Traditional security defense tools include firewall, intrusion detection system, user authentication, and encryption. Firewall through the development of certain access rules to access the request to force inspection only allows the rules of access into the firewall access. Intrusion detection system is another security door behind the firewall. It can monitor the system safely even in the case of fluctuating

network performance and provide real-time protection against internal and external attacks and misoperation. User authentication and encryption, on the one hand, is to ensure the security of data storage and transmission and, on the other hand, is to prevent data leakage.

However, the traditional network security tools mentioned above all start from the security management means and can only passively discover the attack behavior and the existing vulnerabilities after the attack occurs. There is a lag in the security protection, and the defense effect on the discovery of new vulnerabilities and the attack behavior is very small. Attack diagram simulates the whole attack scenario from the attacker's point of view and then presents the dependencies among vulnerabilities in the target network and presents them in a directed graph. According to the attack diagram, the security manager can intuitively observe the vulnerability relationship of each node and select the minimum cost for the security reinforcement of the node that is prone to permission transition. Attack graph is a kind of active network security defense technology, which makes an in-depth analysis of the vulnerability of each host. In the era of frequent network security, attack graph technology has a very good research value [4–7].

The increasing size of the network, the increasing number of security vulnerabilities, and the increasing education of hackers have all made cyberattacks possible. In the game of network security, in order to strengthen the analysis and defense of hacker attack, security defense has gone through many stages of development. At the initial stage, the corresponding attack behavior was matched according to the known attack rules, and then, the vulnerability scanning tool was used to find the vulnerability information on the host computer. However, the method based on rule matching can only find out the single vulnerability information and cannot find out the potential harm caused by the correlation between vulnerabilities. Later, researchers combined network security with statistics and used relevant models to quantify the probability of being attacked. They no longer studied a single category but extended the research target to the whole network system. Then, with the further development of the research, network security researchers have learned that the analysis of network security and other issues should start from multiple dimensions, so a variety of security risk assessment standards and vulnerability analysis methods are proposed. At present, the research hotspot is the security risk assessment method based on the model, which considers the network security from the perspective of the attacker. The model-based research method can show all possible attack paths in the form of graphs and then carry out qualitative or quantitative analysis on the existing security problems, which is convenient for security managers to understand the security of the network environment [8–12]. It is an active network security defense measure. Among all model-based risk assessment studies, the study of attack graph model is the most in-depth. In view of the advantages of attack graph in network security analysis, it has attracted the attention of a large number of researchers, and some difficulties of attack graph have been gradually overcome. However, in the

security analysis based on large-scale network, the large scale of the generated attack graph seriously affects its readability and increases the difficulty in security analysis. Network composite attack modeling is a bit abstract. We analyze the relationship between vulnerabilities in the network from the perspective of attack and draw out the possible attack path of network attack, which can effectively evaluate the direct and indirect security impact caused by network attack. This is to find out the possible attack path, if cooperated with IDS, can be used to predict the attack target. Therefore, now, the study of attack graph automated build technology mainly includes the attack graph and the security analysis of the attack graph. Attack graph automatic construction technology includes target environment modeling technology, vulnerability automatic knowledge base construction technology, attack graph algorithm optimization technology, and attack graph, as shown in Figure 1; the relations between the key technologies of attack graph are as follows. Similarly, the following mainly introduces the research status of attack graph technology from the construction, optimization, and security analysis of attack graph [13–15].

2. Network Security Risk Assessment Based on Attack Graph

2.1. Security Risk Assessment Model. Because network security risk assessment technology plays a positive role in strengthening network defense, researchers have never stopped studying in this field. Initially, researchers had to rely on manual assessments or assistive tools to assess risk. However, manual evaluation is a heavy workload, easy to make mistakes, and has certain subjectivity. Risk assessment tools primarily utilize vulnerability scanning tools. This tool can find information such as server or system environment configuration errors, such as ISS, Nessus, and Nmap. Vulnerability scanning tools mainly use rule matching database to find known vulnerability information. At present, foreign authoritative vulnerability databases include National Vulnerability Database (NVD) of the United States and Bug Trap Vulnerability Database released by Symantec. Authoritative vulnerability databases in China include China National Vulnerability Database (CNNVD). Although the risk assessment tool alleviates the workload of manual assessment to some extent, it can only find known vulnerabilities based on the vulnerability database and cannot find unknown vulnerabilities. Furthermore, only single vulnerabilities can be identified, and neither the correlation between vulnerabilities cannot be found nor the potential harm brought by the correlation between vulnerabilities to the system environment can be evaluated [16–19]. Through RFID, sensing, and other technologies, intelligent bus can understand the location of the bus in real time and realize functions such as turning and route reminder. At the same time, combine with the characteristics of the bus operation, through the intelligent scheduling system, the line, vehicle planning, and scheduling, to achieve intelligent scheduling.

Considering the many defects of manual assessment and auxiliary assessment tools, researchers began to study the relationship between vulnerabilities in system environment

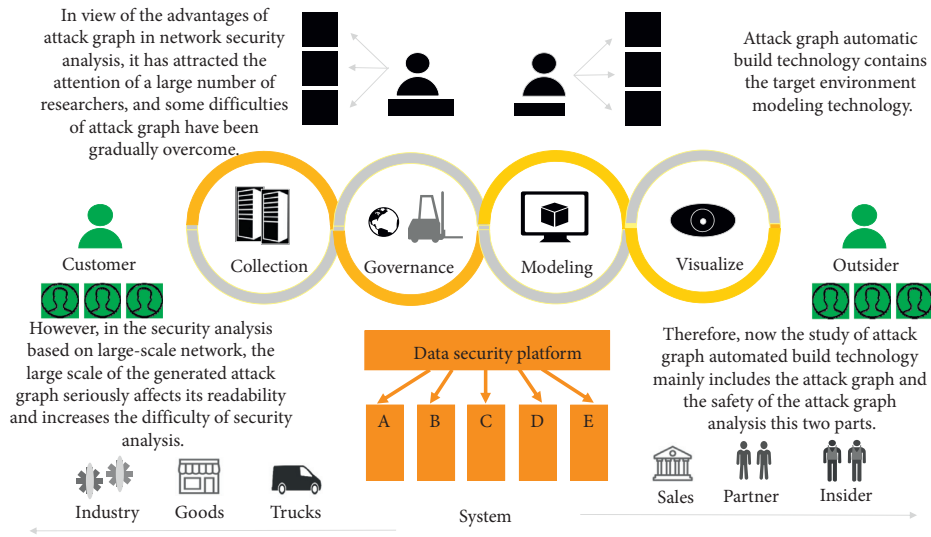


FIGURE 1: Attack graph algorithm optimization technology and attack graph.

by using the security risk assessment method of the model. The model-based security risk assessment method takes the network attack modeling technology as the core, simulates the intrusion scenario of the attacker, and evaluates the network system defense capability qualitatively or quantitatively by means of graph theory. The existing model evaluation methods mainly include fault tree model, attack tree model, privilege graph model, attack graph model, and Bayesian network model. After many years of research on fault tree model, some achievements have been made in determining the deterministic direction of the system. The logical relationship between the fault and the cause of the fault in the fault tree is represented graphically. The logical schematic diagram of the fault tree is shown in Figure 2, where T represents the top event and $S1-S4$ represents the bottom event.

Fault tree is a top-down research method, which analyzes the factors that may cause threats to the network environment and displays all possible failures in the form of graphs. Fault tree can also be called deductive analysis; that is, the analysis starts from the top event (the least expected event) and then from the top of the tree to the leaves of the tree analysis, reverse analysis of the root cause of the top event. The results can be used to analyze how multiple components can cause system failure. The fault tree model includes qualitative analysis and quantitative analysis. Qualitative analysis is to use the fault tree to solve the minimum critical attack set of an event. In quantitative analysis, because the failure rate of a single system event is unknown, only the probability of top event occurrence is calculated [20].

On the basis of fault tree, attack tree targets the inherent threat of the system acquired by different types of attack. Attack tree is also used to describe the whole process of the network system being attacked in a top-down way. The root node in the attack tree represents the attack target, the leaf node represents the attack mode adopted, and the nonleaf node represents the attacked subtarget. The nodes of the

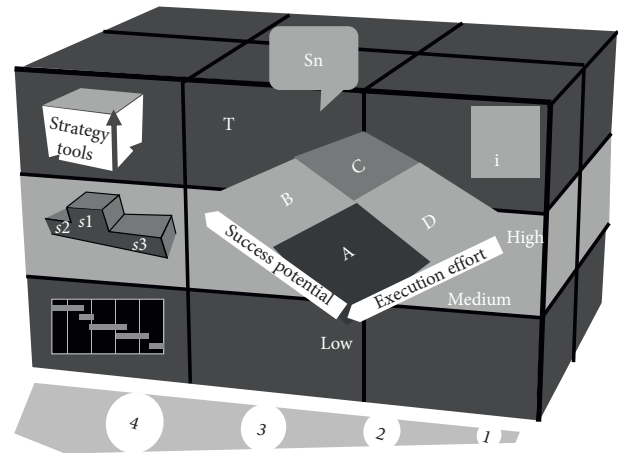


FIGURE 2: The logical schematic diagram of the fault tree.

attack tree fall into two categories: AND nodes and OR nodes. The presentation of AND nodes and OR nodes is shown in Figure 3.

The AND node can only pass the task up when all the child nodes are implemented; the OR node can pass the task up as long as any of the child nodes are implemented.

As proposed in a method for manually building state attack graphs, attack graph model can be regarded as a combination of multiple attack tree models in structure. Attack graph describes the attack behavior of multiple different attack targets and shows the attack path in the way of directed graph. The node and edge in the figure describe the vulnerability or attack behavior of the node, and then according to their logical relationship, the attack path containing the attack starting point, vulnerability node, and attack target is generated. The construction of attack graph model has experienced from manual construction to automatic construction based on model detection and then to the generation of attack graph based on logical reasoning process to adapt to the large-scale network environment,

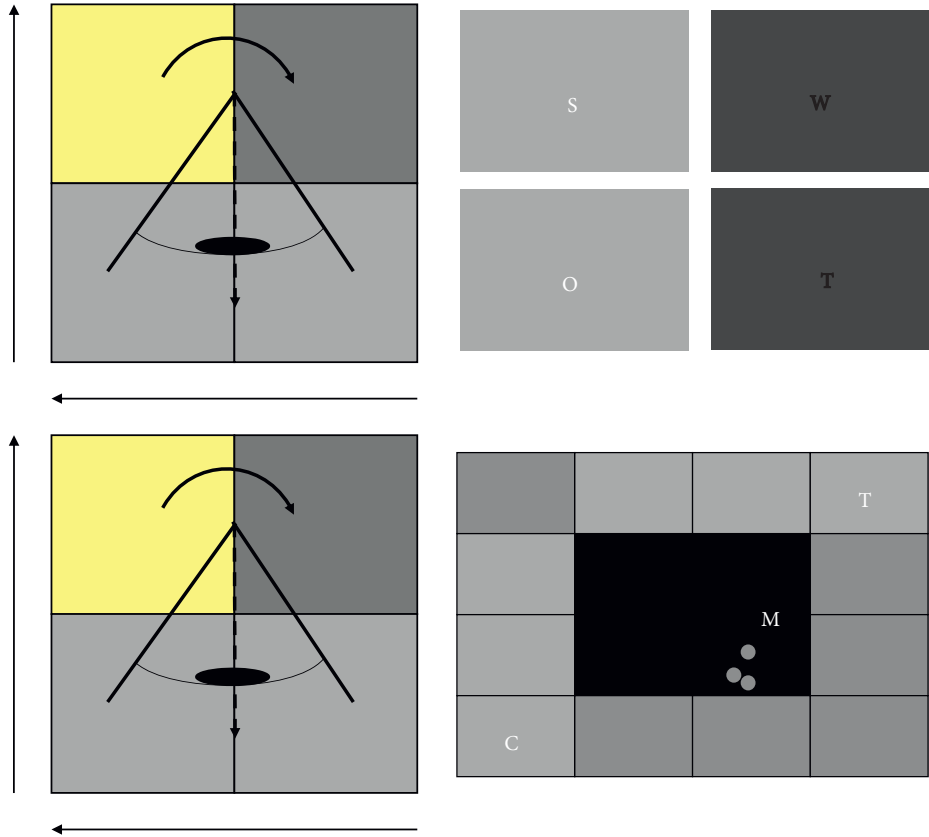


FIGURE 3: The nodes of the attack tree fall into two categories.

thus solving the problem of generating attack graph due to the exponential growth of network scale. Compared with other models, the attack graph model simulates the real attack scenario and detects the vulnerability nodes in the network from the perspective of the attacker. To make the description of network attack more specific and more reliable, the model is as follows:

$$k = \begin{cases} \arg \max \left\{ g(x, y) \eta^\varepsilon(x, y) + \sqrt{\frac{n(n-1)(n+1)}{2}} \right\}, r \neq r_0, \\ \sqrt{\frac{n(n-1)(n+1)}{2}}, \text{ others.} \end{cases} \quad (1)$$

Bayesian network is a model that combines probability analysis and graph theory to describe uncertain causality. Firstly, it assigns probability values to the nodes or edges of the attack graph model and then calculates the attack probability from the initial node to the target node according to the whole attack process and carries out integration processing. The integrated probability is used as the reference index of the whole system network security. Figure 4 is a simple Bayesian network with four variables: the four nodes in the figure correspond to the four real events, and the value of each node is discrete and can only be true or false, so the conditional probability distribution of each node can be described in the form of table.

2.2. Network Security Risk Assessment Process Based on Attack Graph. According to the research status, the attack graph model is the most effective model to represent the dependency and causality of vulnerability and plays an important role in network security assessment. Based on the construction, optimization, and analysis of attack graph-based network security risk, assessment flow chart is proposed in this section. As shown in Figure 5, the network security risk assessment process based on attack graph in this paper is divided into three modules, namely, the formal representation module of abstract network environment, the generation and optimization module of attribute attack graph, and the solution module of key attack set based on attack graph. Abstract network environment formal representation module is not only the basis of network security analysis but also the premise of the latter two modules. This module mainly contains the important file configuration information in the system network. This information can be obtained through vulnerability scanning tools and port scanning tools. The main contents of this paper are the optimization of attribute attack graph generated in large-scale networks and the solution of the minimum critical attack set of simplified attack graph.

Attribute attack graph is constructed with mature tools, which have been studied by predecessors, but the attack graph constructed in large-scale network has some problems such as low readability and unsatisfactory guidance for security managers. To solve this problem, the simplified

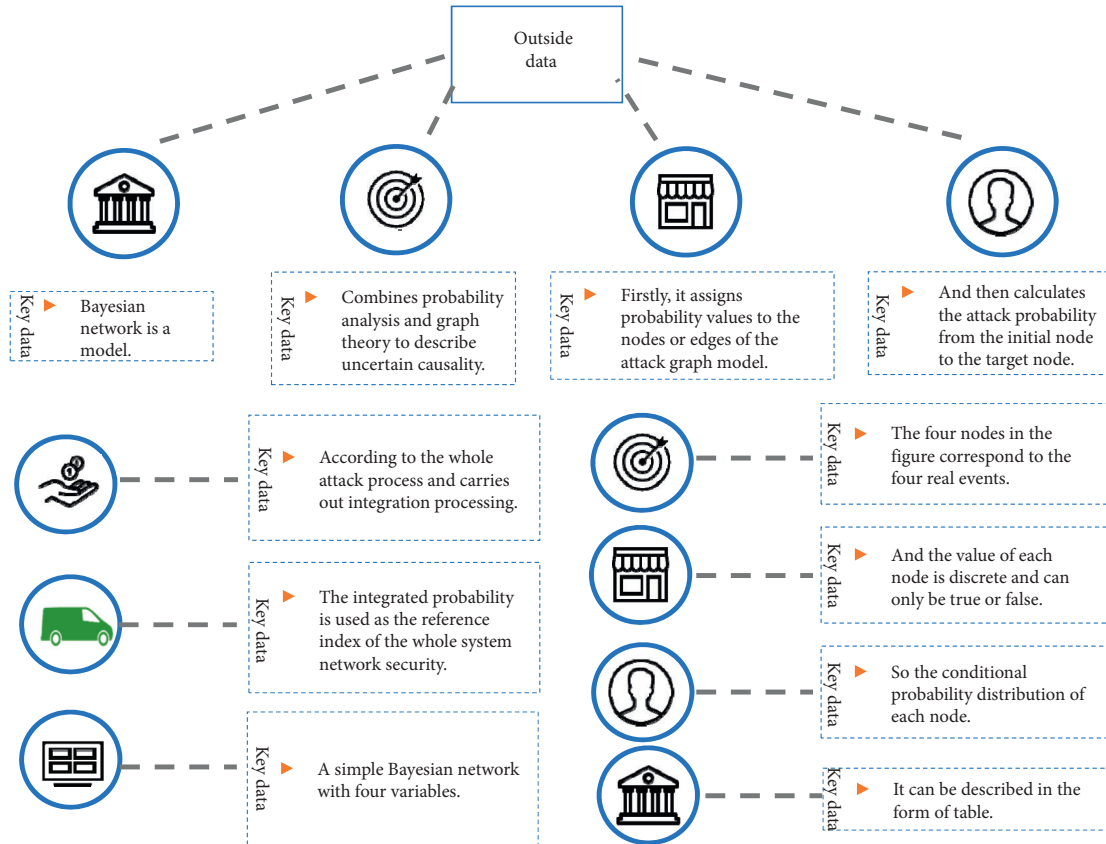


FIGURE 4: A simple Bayesian network with four variables.

attack graph is an improvement in the low readability of the complex attribute attack graph. The attack graph with complex attributes is simplified by using the optimization algorithm based on attack distance and atomic weight, which deletes the meaningless nodes and paths in the real attack path and enhances the readability of the complex

attack graph. Moreover, in the further simplification, the calculation method of path complexity is defined, and attack graphs of different simplified degrees can be obtained by taking different Min value t , which plays a certain role for security managers to predict and judge network risks. The model is as follows:

$$\Omega_{xy}(m+1) = \sqrt{\frac{n(n-1)(n+1)}{2}} + \prod \omega_{xy}^{ck}(m) + (1 - \sigma)^{(m+1/m-1)}, \Omega = \Omega_{\max}, \tag{2}$$

$$\Omega_{xy}(m+1) = \sqrt{\frac{n(n-1)(n+1)}{2}} + \prod \omega_{xy}^{ck}(m) + (1 + \sigma)^{(m+1/m-1)}, \Omega \neq \Omega_{\max}.$$

$$\prod \omega_{xy}^{ck}(m) = \begin{cases} (1 - \sigma)^{(m+1/m-1)}, \Omega = \Omega_{\max}, \\ (1 + \sigma)^{(m+1/m-1)}, \Omega \neq \Omega_{\max}. \end{cases} \tag{3}$$

After the simplified attack graph is obtained, only the key nodes in the attack graph are used for quantitative analysis and security analysis, and the results are uncertain. In the third module, how to solve the minimum critical attack set is proposed, which provides a reliable analysis basis for the minimum network security reinforcement of the network system. Firstly, it is proved that

solving the minimum critical attack set problem is equivalent to the NP-complete problem. In solving the NP problem, the traditional ant colony algorithm will have the problem of premature convergence and slow search speed in the later period, resulting in only obtaining the local optimal solution. Therefore, this paper proposes an improved ant colony algorithm to solve the

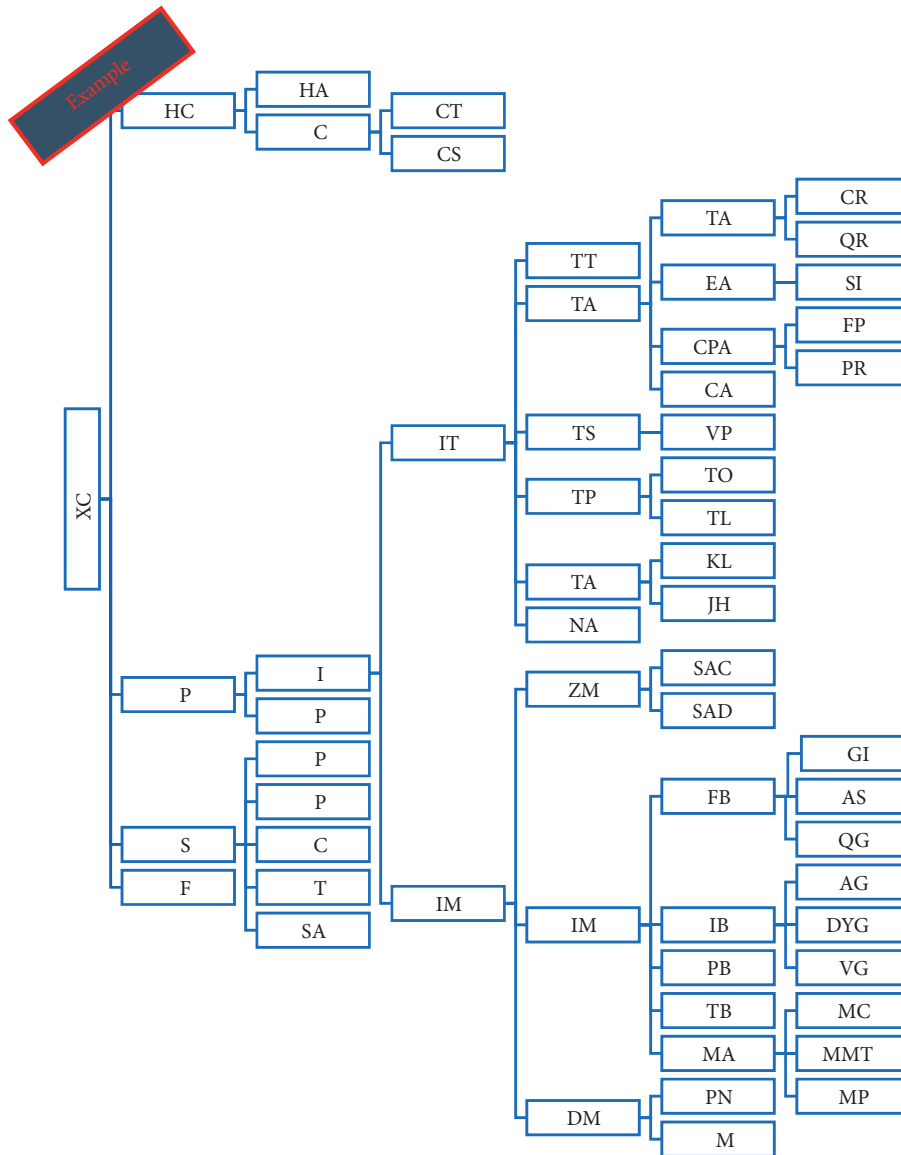


FIGURE 5: The network security risk assessment process based on attack graph.

NP problem. By improving the adaptive update of its pheromone and using the adaptive genetic algorithm to carry out local search, the minimum critical attack set can be solved more accurately and the search speed can be accelerated in the later stage of the algorithm, which provides guidance for the network security management personnel to make the minimum cost network reinforcement measures.

This chapter first introduces several network security risk assessment models and analyzes the advantages and disadvantages of each security risk assessment model in detail. Then, several common attack graph models are introduced, with emphasis on attribute attack graph. On the basis of attribute attack graph, the general formal description and representation of network security elements are given. Finally, this paper describes the security risk assessment flow chart based on the attack graph model and introduces the

function of each module in the flow chart in detail, which lays a foundation for the following research implementation.

3. Analysis and Implementation of Multiobjective Optimization Algorithm of Attack Graph Expert System Model

3.1. *Multitarget Attack Graph Construction Technology.* Attack diagram correlates the vulnerability among hosts in the network, actively discovers the existing vulnerabilities, and represents the possible attack paths in the way of directed graph, but with the expansion of network scale, the problem of space state explosion exists in the construction of attack graph. Too many redundant attack paths in the attack diagram seriously affect the quantitative analysis and accurate judgment of network security by security analysts, thus increasing the cost of network defense. To solve this

problem, this chapter firstly deals with the elimination of the loop in the attack path of the attack graph. Then, in the analysis of the acyclic attack graph, the definition of the path complexity is given based on the attack distance and the weight of atoms. Using different t values to get different degrees of simplified graph, it can improve the readability of the attack graph, which is the possibility of the attack graph being realized in the quantitative analysis of network vulnerabilities.

There are two types of edges in the attribute attack graph: one is the edge from the state node to the attack node, which represents the requirement relationship; the other is the edge from the attack node to the attribute node, which represents the implementation relationship. Figure 6 is an example of a property attack graph. The ellipse represents the attack node, and the text represents the attribute node. When an attacker invades a target network, he first takes advantage of the vulnerability on a host to obtain the initial permissions, and then, he invades again on this basis and repeatedly until the attacker achieves the final target of his attack. Therefore, the process of the attacker's intrusion is actually a transition process from the premise network state to the result network state. Attack graph and related concepts are defined, and construction techniques are explained in the following paragraphs.

The construction of attribute attack graph in this paper adopts the modeling method based on logical reasoning proposed by OU—MuiVal. This modeling method is built on the basis of Prolog logic system and XSB reasoning engine. MuiVal is also an inference system for automatic identification of network vulnerabilities within enterprises. The modeling idea is to first describe the network configuration information and system security policy in a general form, then select its attack behavior as the inference rule, and finally use the XSB inference engine to judge the security policy, if not, then give all possible attack paths. The time complexity of MuiVal is to generate attack graph. MuiVal is a commonly used attack graph building tool. The design principle of MuiVal is to represent the network system configuration information by means of datalog. Attack behavior and defense measures can also be classified using datalog.

Attribute attack graph has two nodes: attribute and attack, which show attack path implicitly. Therefore, there may be multiple attack nodes repeatedly attacking an attribute node, but there is also a relationship between attack nodes with jumping permissions, so it is easy to cause the phenomenon of attack graph loop. The existence of loop makes the relationship between nodes more complex, which has a serious impact on the accuracy of attack probability calculation and the readability of attack graph. It is found that the loop phenomena of attack graph can be divided into three categories.

As shown in Figure 7, text represents the attribute node and ellipse represents the attack node.

But in the multitarget network, deleting a node in the third type of loop may delete the attack path with high-risk probability. In the further study, it is found that if the forward search or depth-first search method is used to calculate all the reachable parent nodes. In the iteration process, in order to eliminate the ring of attack path, the

trace set was defined to place the tracked attribute node that was searched. Before iteration, place all the child nodes of the attribute node to trace the collection and then iterate the child nodes of the node in attack, and in the process of iteration, if the attribute node of the child node is found to be placed in the track, or the child node is found to have been tracked, then the iteration will continue to enter the loop until the iteration is terminated.

As an intelligent subject, the attacker often chooses the attack path with short attack distance and low attack complexity to attack when carrying out network attack. Based on this fact, there are a lot of redundant attack paths in the attack graph. Therefore, in the optimization process of attack graph, if the path complexity of each attack path is calculated according to the attack distance and atomic weight and then a value is set to remove those lower than these safe paths, the attack graph will be simplified to a large extent. Calculation of path complexity is as follows:

$$R = \begin{cases} \prod \omega_{xy}^{ck}(m), m \triangleq 125, \\ \sqrt{\frac{n(n-1)(n+1)}{2}}, \text{ others.} \end{cases} \quad (4)$$

In order to verify the feasibility of eliminating loops for these three types of loops and the scalability of path complexity, we conducted two experiments. In the first part of the experiment, we use the MuiVal attack graph building tool to generate attack graph for a certain scale network environment, analyze the loop condition of attribute attack graph, give the idea of eliminating the loop, and get the acyclic attack graph. The second part of the experiment is to calculate the path complexity of acyclic attack graph, according to different values can get different scales of simplified attack graph. For different values, if different thresholds are set, the number of paths reserved in the simplified graph can be obtained is different. The result is shown in Figure 8.

It can be concluded from Figure 8 that the number of attack paths reserved under different Min values is also different. As the selection of threshold increases, the number of paths reserved decreases. When the value reaches a certain value, all possible attack paths must be removed. Therefore, the selection of the value is also empirical to a certain extent. If the value is too high, the simplified graph is too simple, and it is easy to delete the important nodes of the original attack graph, resulting in the wrong defense reinforcement of the attack nodes. If the value is selected too low, the simplified graph is still complicated, which makes the visibility of attack graph low and the risk assessment for security managers still very difficult.

4. Attack Diagram Expert System Model Multitarget Network Security Algorithm Evaluation

Genetic algorithm is a random search method imitating the biological evolution process. By sacrificing the limited iteration time to optimize the search speed, the algorithm can

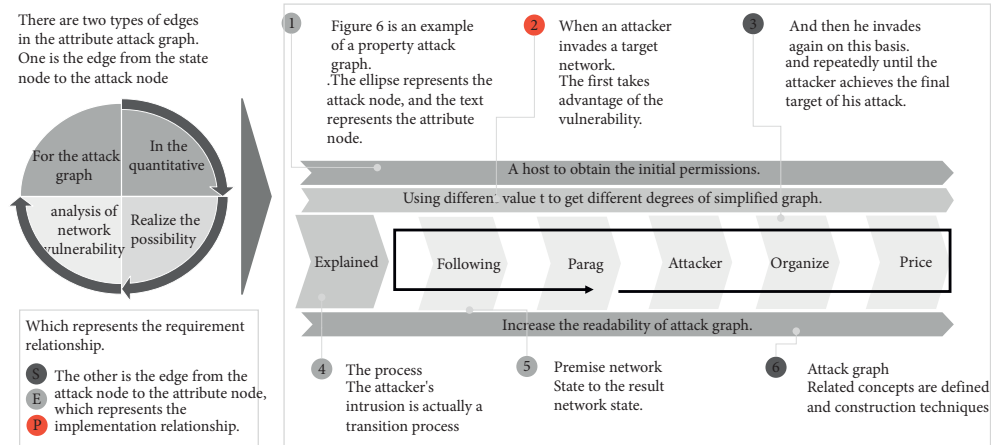


FIGURE 6: An example of a property attack graph.

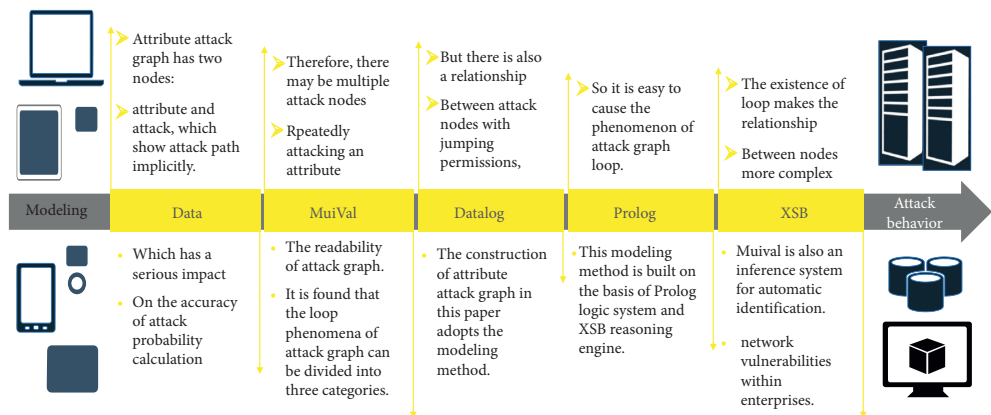


FIGURE 7: Text represents the attribute node and ellipse represents the attack node.

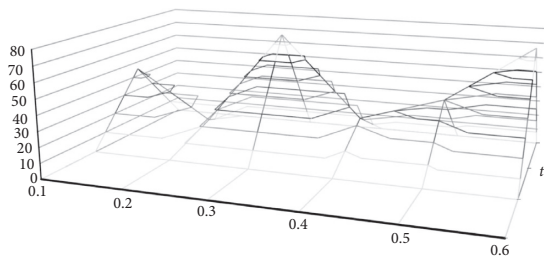


FIGURE 8: The number of paths reserved in the simplified graph.

automatically obtain and guide the optimal search space and adjust the search process adaptively. Because of its bionic nature, it can provide another way to solve problems that other science and technology cannot or are difficult to solve. It has been widely used in combinatorial optimization, machine learning, artificial intelligence, and other fields. However, the traditional genetic algorithm has strong robustness in solving the global optimal solution, which leads to some characteristics such as premature convergence and slow search in the late period. To solve this problem, an adaptive genetic algorithm was proposed, which made the ant colony self-adjust the crossover rate and mutation rate according to its own race. The optimization problem can be

described by the following mathematical programming model.

Figure 9 describes the variation trend of the fitness of the three algorithms. It can be seen from the figure that the adaptive genetic algorithm maintained a high search speed in the early and middle periods and began to slow down its search speed in the 160–180 generation. In the 30th generation, genetic algorithm and adaptive sorting algorithm began to show a step-down state. In the 60th generation, their search speed was slow, and even in the 20th generation, their search speed was almost stagnant. As can be seen from the above, when different values of t are taken, the simplified graphs obtained are also different, and the minimum critical attack sets obtained through the improved ant colony algorithm are also different. However, the security manager can select the values according to the security management resources he/she has, so as to ensure the minimum security network reinforcement measures.

Figure 10 describes the variation curve of the performance of the adaptive genetic algorithm between the optimal individual value and the population mean value. It can be seen from the figure that the convergence performance of the adaptive genetic algorithm is relatively slow in the 0–80 generation and gradually begins to converge to complete the solution of the optimal solution after the 80 generation.

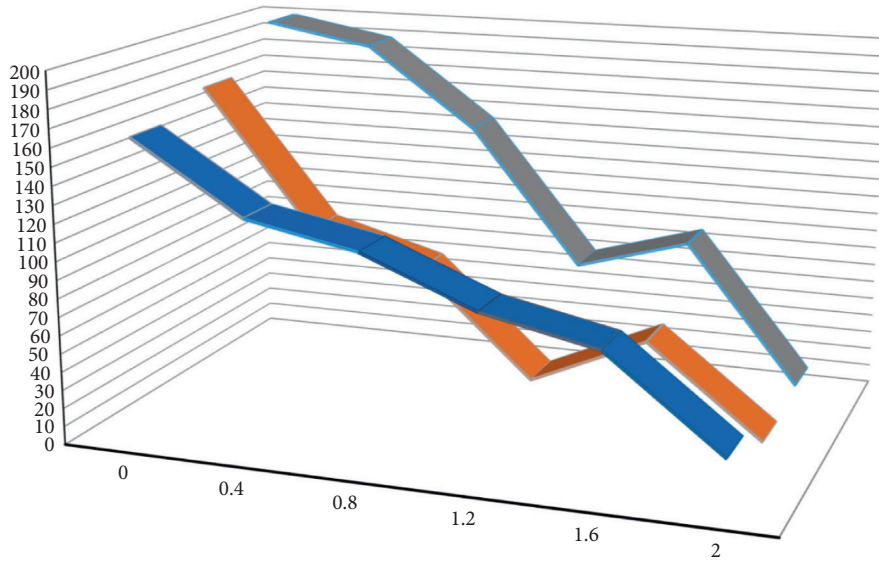


FIGURE 9: The variation trend of the fitness of the three algorithms.

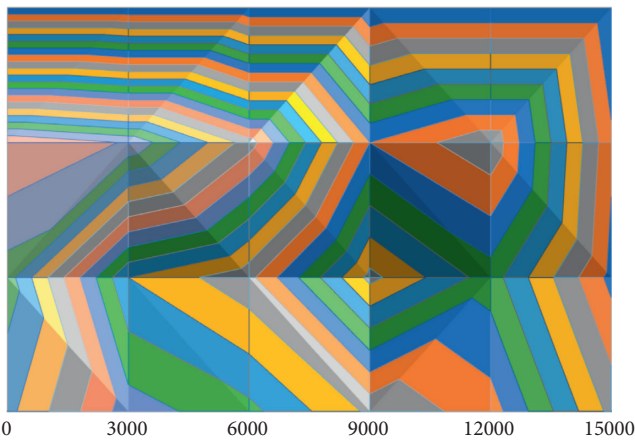


FIGURE 10: The variation curve of the performance of the adaptive genetic algorithm.

It can be seen from Figure 11 that, under the same threshold value t , the minimum critical attack set can be solved for attack graphs of different scales. It can be seen that the improved ant colony algorithm can solve the key attack set more optimally. In addition, it can be seen from the figure that, in the case of increasing attack paths in the attack graph, the optimal solution effect of the improved ant colony algorithm is more obvious and closer to the optimal solution. At the same time, the search speed of the improved ant colony algorithm is also further faster than that of the ant colony algorithm to some extent. For large-scale network, the search speed of the improved ant colony algorithm will not slow down in the later period. In the above five scale attack graphs, the search speed of the improved ant colony algorithm is 8.62% higher than that of the ant colony algorithm on average, and the highest efficiency is three times than that of other methods. Also, it is found that, in the case of many attack paths, the improved ant colony algorithm improves the search speed optimization more obviously and is more suitable for large-scale network.

In this section, the concept of minimum critical attack set (MCPS) is first expounded, and then, it is proved that the solution of MCPS can be equivalent to the solution of NP-complete problem. Ant colony algorithm (ACO) is widely used in solving NP complete problems. However, it is found in the further solution that ACO tends to converge prematurely in the solution process, which often leads to convergence at the local optimal solution and makes the algorithm stagnate instead of getting the global optimal solution. Therefore, this paper proposes an improved ant colony algorithm in the fourth section, which uses the adaptive update of pheromone to optimize the selection of the next node and uses the adaptive genetic algorithm to improve the local search strategy. In the final experimental analysis, the results show that the improved ant colony algorithm has a significant improvement in solving the problem of the minimum critical attack set and in searching speed under different network sizes. Especially when the scale of attack graph increases gradually, the effect of solving accuracy and solving time becomes more obvious, and it is more suitable for large-scale network. Most of the time vulnerabilities and their disclosure are due to poor system management, late patching, weak password policies, inadequate access control mechanisms, and so on. Therefore, the primary reason and purpose of penetration testing should be to identify and correct failures in the system management process that lead to the emergence of system vulnerabilities that are disclosed during penetration testing. The method presented in this paper can effectively avoid such leakage.

Many cities begin to introduce intelligent transportation facilities into the urban transportation construction, which can not only guarantee people's travel more smoothly but also further promote the development and progress of the transportation field. Although intelligent transportation has begun to be effectively applied to cities, as the urban population surges, many rural people go to cities to work, making the development of

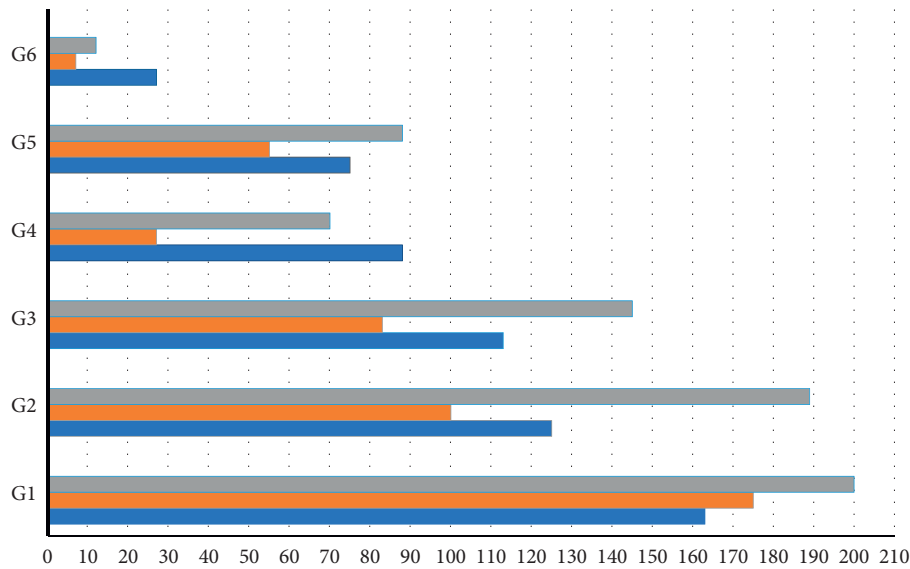


FIGURE 11: The minimum critical attack set can be solved for attack graphs of different scales.

intelligent transportation difficult to withstand the ever-expanding traffic pressure in cities. It can be found that the traffic conditions of many cities in our country are very tight, and the traffic pressure problem is increasingly aggravating. In order to make the intelligent transportation system better meet the needs of the city, the transportation system and the Internet of Things should be organically combined, so that the Internet of Things technology can be applied scientifically and the transportation facilities can realize the joint application of the intelligent transportation system and solve the problems faced by the urban transportation.

5. Conclusion

Nowadays, computer technology has entered our daily life, so it is urgent to protect and maintain network security. However, in the face of the diversity and complexity of the attack means of the attackers on the network, our security protection measures have a certain lag. In the final analysis, the problem of network security is due to the existence of vulnerability loopholes in the network system, which gives attackers an opportunity to take advantage of. Therefore, in order to be able to develop better preventive measures in the network security protection system, we should start from the analysis of the vulnerability vulnerabilities in the network system. Therefore, this paper focuses on the network security evaluation and analysis based on the attack graph model. Attack graph technology can analyze the vulnerability of each host in the network, find out the attack path threatening network security, and show it in the way of directed graph. However, the attack graph generated by large-scale network is too complex and poor in readability, which is difficult for security managers to analyze, thus affecting their accurate security judgment and reinforcement of defense measures. Therefore, the main work of this paper is as follows: using MuiVal attack graph building tool to generate a directed

attribute attack graph. The phenomena of various kinds of loops in attribute attack graph are explained in detail, and the general methods of eliminating loops are expounded. An optimization algorithm based on path complexity is implemented for the attack graph after eliminating the loop. The attack graph can be optimized to different degrees by using the formula of attack distance and atomic weight and setting different values. Finally, the effectiveness of the proposed algorithm is verified by an experiment based on a specific network topology. On the basis of simplifying the attack graph, the network environment is strengthened with the minimum cost. Firstly, the concept of minimum critical attack set (MCCS) is expounded, and then, it is proved that solving the MCCS is equivalent to solving the NP problem. According to ant colony algorithm, the problem of early convergence and late search is too slow in solving the minimum critical attack set so that only the local optimal solution can be obtained. Based on ant colony algorithm, this paper proposes to improve ant colony algorithm by using adaptive updating principle of pheromone and local search strategy. Finally, the simulation experiment is used to compare the two algorithms in the approximate solution speed and the number of sets.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] M. H. R. Khoisan, L. Zhenjiang, and M. Pasquale, "Scalable min-max multi-objective cyber-security optimisation over probabilistic attack graphs," *European Journal of Operational Research*, vol. 278, no. 3, pp. 894–903, 2019.
- [2] H. Hu, J. Tan, and J. Liu, "SOCMTD: selecting optimal countermeasure for moving target defense using dynamic game," *KSI Transactions on Internet and Information Systems (TIIS)*, vol. 14, no. 10, pp. 4157–4175, 2020.
- [3] M. Stolpe, "The internet of things," *ACM SIGKDD Explorations Newsletter*, vol. 18, no. 1, pp. 15–34, 2016.
- [4] E. D. Pontiki, "Multi-target cinnamic acids for oxidative stress and inflammation: design, synthesis, biological evaluation and modeling studies," *Molecules*, vol. 24, no. 1, p. 12, 2019.
- [5] A. O. Akmandor, H. Yin, and N. K. Jha, "Smart, secure, yet energy-efficient, Internet-of-Things sensors," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 4, no. 4, pp. 914–930, 2018.
- [6] G. Spanos and L. Angelis, "A multi-target approach to estimate software vulnerability characteristics and severity scores," *Journal of Systems and Software*, vol. 146, no. 7, pp. 152–166, 2018.
- [7] L. Yuqing, "An intelligence-driven security-aware defense mechanism for advanced persistent threats," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 646–661, 2018.
- [8] M. O. Al Kalaa and H. H. Refai, "Monitoring radiated co-existence testing using GMM-based classifier," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10336–10345, 2017.
- [9] S. Thöns, M. G. Stewart, and G. S. Mark, "On decision optimality of terrorism risk mitigation measures for iconic bridges," *Reliability Engineering & System Safety*, vol. 188, no. 13, pp. 574–583, 2019.
- [10] P. Lau, W. Wei, L. Wang, Z. Liu, and C.-W. Ten, "A cybersecurity insurance model for power system reliability considering optimal defense resource allocation," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4403–4414, 2020.
- [11] A. R. Soury, "A state-of-the-art survey of malware detection approaches using data mining techniques," *Human-centric Computing and Information Sciences*, vol. 8, no. 1, pp. 1–22, 2018.
- [12] A. R. Kumar and A. Sivagami, "Security aware multipath routing protocol for WMSNs for minimizing effect of compromising attacks," *Journal of Network and Systems Management*, vol. 27, no. 3, pp. 573–599, 2019.
- [13] S. Velliangiri, "A hybrid BGWO with KPCA for intrusion detection," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 32, no. 1, pp. 165–180, 2020.
- [14] G. Tao and Q. Zhang, "Research on smart grid energy internet of things technology and network security," *IOP Conference Series: Materials Ence and Engineering*, vol. 677, no. 4, pp. 42–117, 2019.
- [15] J. Wang, A. Liu, and T. Yan, "A resource allocation model based on double-sided combinational auctions for transparent computing," *Peer-to-Peer Networking and Applications*, vol. 11, no. 4, pp. 612–633, 2018.
- [16] X. Feng, H. Zhang, and Y. Ding, "A review study on traction energy saving of rail transport," *Discrete Dynamics in Nature and Society*, vol. 2013, no. 1, pp. 1611–1617, 2013.
- [17] V. Miori, D. Russo, and L. Ferrucci, "Supporting active aging through a home automation infrastructure for social internet of things," *Advances in Science, Technology and Engineering Systems Journal*, vol. 3, no. 4, pp. 173–186, 2018.
- [18] A. Aliverti, "Wearable technology: role in respiratory health and disease," *Breathe*, vol. 13, no. 2, pp. e27–e36, 2017.
- [19] A. Manocha, R. Singh, and P. Verma, "An internet of things fog-assisted sleep-deprivation prediction framework for spinal cord injury patients," *Computer*, vol. 53, no. 2, pp. 46–56, 2020.
- [20] M. S. Mahmud, H. Fang, and H. Wang, "An integrated wearable sensor for unobtrusive continuous measurement of autonomic nervous system," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 1104–1113, 2018.

Research Article

Deep Multiple Kernel Learning for Prediction of MicroRNA Precursors

Hengyue Shi , Dong Wang, Peng Wu , Yi Cao, and Yuehui Chen

Information Science and Technology, University of Jinan, Jinan 250022, China

Correspondence should be addressed to Peng Wu; ise_wup@ujn.edu.cn

Received 12 March 2021; Revised 1 April 2021; Accepted 19 April 2021; Published 5 May 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Hengyue Shi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

MicroRNAs are a group of noncoding RNAs that are about 20–24 nucleotides in length. They are involved in the physiological processes of many diseases and regulate transcriptional and post-transcriptional gene expression. Therefore, the prediction of microRNAs is of great significance for basic biological research and disease treatment. MicroRNA precursors are the necessary stage of microRNA formation. RBF kernel support vector machines (RBF-SVMs) and shallow multiple kernel support vector machines (MK-SVMs) are often used in microRNA precursors prediction. However, the RBF-SVMs could not represent the richer sample features, and the MK-SVMs just use a simply convex combination of few base kernels. This paper proposed a localized multiple kernel learning model with a nonlinear synthetic kernel (LMKL-D). The nonlinear synthetic kernel was trained by a three-layer deep multiple kernel learning model. The LMKL-D model was tested on 2241 pre-microRNAs and 8494 pseudo hairpin sequences. The experiments showed that the LMKL-D model achieved 93.06% sensitivity, 99.27% specificity, and 98.03% accuracy on the test set. The results showed that the LMKL-D model can increase the complexity of kernels and better predict microRNA precursors. Our LMKL-D model can better predict microRNA precursors compared with the existing methods in specificity and accuracy. The LMKL-D model provides a reference for further validation of potential microRNA precursors.

1. Introduction

MicroRNAs are a class of highly conserved endogenous noncoding RNAs with a length of about 20–24 nucleotides. They are single stranded and regulate gene expression at the post-transcriptional or translational level by binding specifically to target messenger RNA [1, 2]. Studies have shown that some microRNAs can play a role by regulating cell proliferation, cell migration, invasion, and immune response [3], and at the same time, microRNAs can also play an important role in inflammatory response [4], neural development, and other processes [5, 6]. In organisms, microRNA is first transcribed by RNA polymerase II into long initial transcription, primary microRNA, which is then processed by Drosia enzyme into a precursor with a length of about 70 nucleotides, that is, pre-microRNA [3, 7]. Pre-microRNA is exported from the nucleus with the help of RanGTP/exportin 5 and then exported to be processed and matured by the Dicer enzyme in the cytoplasm [8, 9]. After

being processed into mature microRNAs, microRNAs form RNA-induced silencing complex (RISC) in some way to affect the protein abundance of target genes by inhibiting translation or degrading the mRNAs of target genes.

MicroRNA precursors (pre-microRNAs) can fold into hairpin structures, which are considered the most important indicators of microRNA maturation [10]. Figure 1 shows a pre-microRNA sequence and its hairpin structure. However, there are a large number of nonprecursors with similar hairpin structures in many genomic regions, which are called pseudo hairpin sequences [11]. Accurately and effectively identifying microRNA precursors from a large number of candidate hairpin sequences is a challenging task [12].

The methods of finding new microRNAs mainly include biological experimental methods and computer prediction methods [13]. Biological experimental methods are more direct and highly reliable, but the expression level of microRNAs is relatively low. Some microRNAs are only

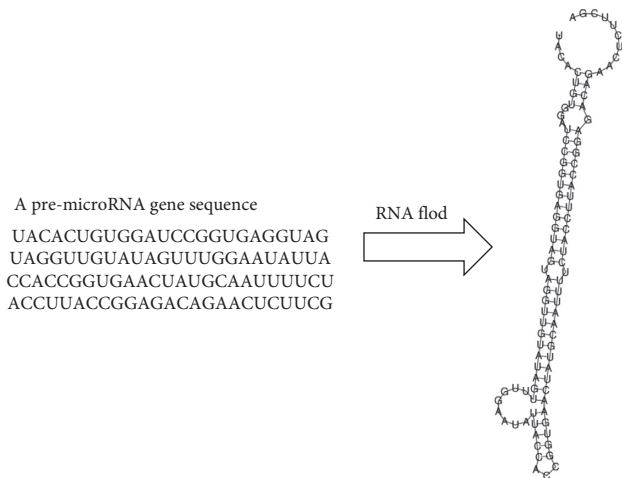


FIGURE 1: Structure of the pre-microRNA.

expressed under specific conditions, such as cell type and physiological state of the body. Moreover, due to the high cost and long experimental cycle, it is difficult to replicate microRNA expressed in a specific tissue and period. With the help of computer, the computer prediction method can identify new microRNAs more efficiently. The prediction method of microRNAs based on machine learning has been applied to bioinformatics, which can overcome various defects of biological experimental methods, prevent microRNAs from being affected by expression time, tissue specificity, or expression level, and provide reliable samples for subsequent biological experiments.

MicroRNA precursors have a unique hairpin structure and are easier to obtain than microRNAs. Thus, computational prediction methods use machine learning to mainly identify microRNA precursors from candidate hairpin sequences. The authors in [14] and [15] proposed a set of novel features and used a support vector machine (SVM) with only a RBF kernel to classify real and pseudo pre-microRNAs and proposed Triplet-SVM and PremipreD. While different kernels have different characteristics, a RBF kernel could not adequately map the pre-microRNAs to the appropriate feature spaces. When the features of input data contain heterogeneous information [16, 17] or the data are nonflat in the high-dimensional feature space [18], it is not reasonable to use a single simple kernel to map all the input data. The authors in [19] used a random forest classifier to find shallow features and only got 91.29% accuracy and proposed MipIe. The authors in [20] adopted multiple kernel SVM with different weights, but only the shallow features are used and then LMKL-MiPred was proposed. It shows good accuracy but no deep features of pre-microRNAs were explored. The authors in [21] used a simple three-layer backpropagation neural network and proposed MiRANN. However, when there are limited candidate hairpin sequences, the three-layer backpropagation neural networks typically do not have a good generalization performance, and they can even increase the risk of over-fitting under some conditions.

Multiple kernel methods have been successful on small data sets. By mapping the samples into a high-dimensional

reproducing kernel Hilbert space, they only use very few parameters to enable a classifier to learn a complex decision boundary. How to determine the basic kernel function is the difficulty and key problem of multiple kernel learning. The localized multiple kernel learning [22] uses different weights to combine simple basic kernel (linear kernel, polynomial kernel, and RBF kernel) but could not obtain the deep features of the samples. This paper presents a localized multiple kernel learning model with a nonlinear deep synthetic kernel (LMKL-D). The deep synthetic kernel was trained by a deep multiple kernel learning model with a tree structure [23]. We found that the neural networks are easy to obtain deep features by gradient descent. Thus, we adopt the gradient descent approach and use a deep multiple kernel learning model with a tree structure to get a nonlinear deep synthetic kernel. We combine kernels at each layer and then optimize over an estimate of the leave one out error [23]. Starting from some simple basic kernels, a deep synthetic kernel can be achieved after a learning process. We combined the deep synthetic kernel and other simple basic kernels by localized multiple kernel learning. The deep synthetic kernel was composed of complex combination of basic kernels. Thus, the LMKL-D model can take advantage of both the shallow and deep features of the input data. As a result, the LMKL-D model can represent more features and obtain better performance than existing SVM methods.

The rest of the paper is organized as follows. In Section 2, we introduce datasets selection and feature selection and then we provide the background about SVM, localized multiple kernel learning, and the multiple kernel learning methods. Kernels and model selection are also included in this section. In Section 3, we show the experimental results and comparisons with other methods. Finally, conclusions and future work are shown in Section 4.

2. Materials and Methods

2.1. Biologically Relevant Datasets. The LMKL-D model proposed in this paper should be able to correctly identify pre-microRNAs and pseudo hairpin sequences from the candidate hairpin sequences dataset. Thus, the candidate hairpin sequence datasets have two parts. One is the positive real pre-microRNAs sequences. We obtained a total of 4,028 annotated known pre-microRNA sequences spanned 45 species from miRBase 12 [24]. We removed sequences with homology greater than 90% from the original sequences, and finally 2,241 nonhomologous pre-microRNAs were selected as positive sequences. The other part is the negative pseudo hairpin sequences. The pseudo hairpin sequences were obtained from the UCSC refGene annotation list [25] and the human RefSeq gene [26]. For pseudo hairpin sequences, their sequence fragments have similar hairpin structures of pre-microRNAs and were not reported as pre-microRNAs. Finally, we selected 8,494 pseudo hairpin sequences from the protein coding region. These sequences must be guaranteed to be around 90 ribonucleotides, with a minimum of -15 kcal/mol and a maximum 18 kcal/mol free energy.

In order to select better model, we randomly selected seventy percent of the candidate hairpin sequences as the

training set and the remaining thirty percent as the test set. Thus, we randomly selected 1,500 pre-microRNAs and 6,000 pseudo pre-microRNAs as the training set. As for test set, 700 of the remaining positive real pre-microRNAs and 2,400 of the remaining negative pseudo hairpin sequences were randomly selected. Both training set and test set are normalized.

2.2. Feature Selection. There are many methods to select the pre-microRNAs features. Traditionally, sequence, secondary structure, and thermodynamic properties are considered. In this paper, we use the dinucleotide frequencies proposed in [27] to characterize sequence and secondary structure properties. Thermodynamic characteristics are also included. The LMKL-D model assumed that the hairpin structure of each sequence could be individually characterized as an eigenvector containing 29 global and intrinsic folding properties [27, 28]. Seventeen attributes are the A; C; G; U dinucleotide frequencies; and G + C ratio; three attributes are the folding measures, including the base pairing propensity, base pair distance, and Shannon entropy; three thermodynamic properties such as minimum free energy (MFE) of folding, MFE index 1 MFEI₁, and index 2 MFEI₂; one topological attribute; and five normalized variants of folding measures. The sequence properties and thermodynamic properties can be calculated by ViennaRNA Package 2.0 [29].

2.3. Kernels and Support Vector Machine. The kernels are the inner products of the mapping relationship. A kernel can be described by the dot product of its two basic mapping functions as follows [23]:

$$K(x, y) = \phi(x) \cdot \phi(y), \quad (1)$$

where $K(x, y)$ represents a kernel and $\phi(x)$ and $\phi(y)$ represent the mapping functions.

The mapping functions $\phi(x)$ and $\phi(y)$ are hard to find, but the dot product of the two mapping functions can be easily calculated by the kernel matrix [30]. We can use the characteristics of the kernels to construct a new kernel that can enhance the ability to represent richer features. Thus, the new kernel can map the input data from the low-dimensional linear indivisible feature space to a high-dimensional linearly separable feature space. Synthetic kernels can create different representations of the data using basic kernels.

Kernels are usually associated with SVMs. The basic principle of a single kernel SVM is, for a given dataset $x_i \in R^n$ ($i = 1, \dots, l$) with corresponding labels y_i ($y_i = +1$ or -1), SVM finds the linear separable hyperplane with the maximum margin in the feature space induced by the mapping functions $\phi(x)$ and $\phi(y)$. Equation (2) is the SVM decision function [31] given as follows:

$$f(x) = \sum_{i=1}^l \alpha_i y_i K_\theta(x_i, x) + b, \quad (2)$$

where α_i are the coefficients to be learned and $K_\theta(x_i, x)$ is a kernel that depends on a set of parameters θ . Traditionally,

parameters α_i are trained through maximizing the dual objective function as the following equation [31]:

$$\max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j K_\theta(x_i, x_j) \\ \alpha_i \geq 0, \quad i = 1, \dots, l \quad (3)$$

$$\text{s.t.} \quad \sum_{i=1}^l \alpha_i y_i = 0.$$

2.4. Multiple Kernel Learning. Multiple kernel learning model is a kind of kernel-based learning model with more flexibility. Recent theories and applications have proved that using multiple kernels instead of single kernel can enhance the interpretability of the decision function and obtain better performance than the single kernel model [31, 32]. Multiple kernels can create different representations of the input data using basic kernels. When we combine multiple kernels within a kernel such as by taking their sum, we can obtain a new kernel that is different from each of them. Moreover, the new kernel has more complicated representations that could not be well represented by a single kernel [33, 34].

In the multiple kernel learning model, K_θ is considered as a linear convex combination of multiple basis kernels [31]:

$$K_\theta = \sum_{i=1}^m \theta_i K_i,$$

$$\text{s.t.} \quad \theta_i \geq 0, \quad i = 1, \dots, m, \quad (4)$$

$$\sum_{i=1}^m \theta_i = 1,$$

where K_i are the basic kernels and m is the total number of basic kernels.

2.5. Deep Multiple Kernel Learning. The traditional multiple kernel learning method is just a simple linear combination of a set of basic kernels and could not represent the deep features of the samples. Thus, we adopt a three-layer multiple kernel learning model to represent the deep features of the samples [23]. The complex combination of basic kernels still meets Mercer standards. A deep multiple kernel model is a n -layer multiple kernel model with m kernels at each layer:

$$K^{(n)} = \left\{ \theta_1^{(n)} K_1^{(n)} \left(\theta_1^{(n-1)} K_1^{(n-1)} + \dots \right) + \dots + \theta_m^{(n)} K_m^{(n)} (\dots) \right\}, \quad (5)$$

where $K_m^{(n)}$ represents the m^{th} kernel at layer n with an associated weight parameter $\theta_m^{(n)}$ and $K^{(n)}$ represents the synthetic kernel at layer n . A deep multiple kernel learning model with n layers is shown in Figure 2.

Although the increased complexity of the kernels can increase the risk of over-fitting, Strobl et al. [23] proved that the upper bound of the generalization error for deep multiple kernels is significantly less than that for deep feedforward neural networks under some conditions.

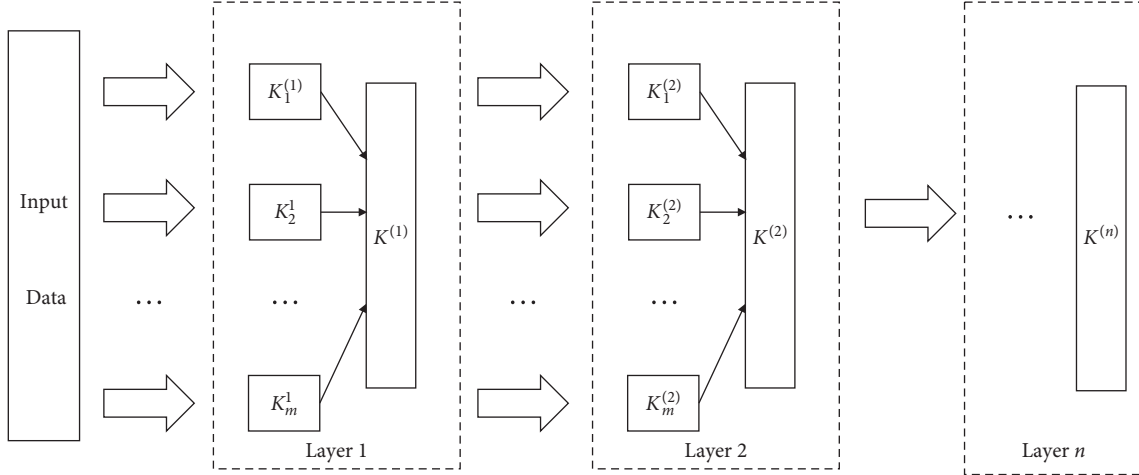


FIGURE 2: A deep multiple kernel model with n layers in this paper. The black solid lines represent the weights for each kernel, $\theta_m^{(n)}$. $K^{(n)}$ comes from the sum of each $K_m^{(n)}$ multiplied by its associated weight $\theta_m^{(n)}$ at layer n .

Leave one out error has shown better accuracy in multiple kernel learning [31]. To decide the weight parameter $\theta_m^{(n)}$, we adopted concept of the span of support vectors [35]. The main idea of SVMs is that mapping input data into a high-dimensional feature space where a hyperplane can separate the input data. The hyperplane can be constructed by maximizing the margin. It is well known that the error rate for SVM is bounded by $\sqrt{O(R^2/M^2)}$, where R is the radius of the smallest ball containing the training data in the feature space and M is the margin. The smaller the error, the better the SVM performance. However, traditional SVM methods only maximize M while R may still be very large. To address the above problem, we minimize R based on upper bounds of leave one out error. R can be shown in span. The span bound of the leave one out error can be shown as the following equation [31]:

$$L((x_1, y_1), \dots, (x_l, y_l)) \leq \sum_{p=1}^l \phi(\alpha_p^0 S_p^2 - 1) =: T_{\text{span}}, \quad (6)$$

where L is leave one out error and S_p is the distance between the point $\phi_{K_\theta}(x_p)$ and the set Γ_p . Γ_p is the linear combination of support vectors mapping into feature space, $\Gamma_p = \left\{ \sum_{i \neq p, \alpha_i^0 > 0} \lambda_i \phi_{K_\theta}(x_i) \mid \sum_{i \neq p} \lambda_i = 1 \right\}$.

We use a contracting function $\phi(x) = (1 + \exp(-cx + d))^{-1}$ to smooth equation (6) and evaluate its performance. Here, c and d are nonnegative arguments and then a regularization term is added to prevent over-fitting [23]:

$$\overline{S_p^2} = \min_{\lambda, \sum \lambda_i = 1} \left\| \phi_{K_\theta}(x_p) - \eta \sum_{i \neq p} \lambda_i \phi_{K_\theta}(x_i) \right\|^2 + \eta \sum_{i \neq p} \frac{1}{\alpha_i} \lambda_i^2. \quad (7)$$

Span is optimized using the gradient descent method. Now, we get the deep multiple kernel learning algorithm with the derivative of $(\partial T_{\text{span}} / \partial \theta)$. By using gradient descent, θ and α can be solved by fixing θ and solving for α and fixing α and solving for θ .

In the deep synthetic kernel proposed in this paper, the number of layers was set as three layers and each layer was set as 3 kernel functions. The kernel functions of the first layer were linear kernel, polynomial kernel, and Gaussian kernel.

2.6. Localized Multiple Kernel Learning. While a single kernel function has only one characteristic, multiple kernel learning (MKL) has more flexibility by choosing a combination of basic kernels. However, multiple kernel learning assigns the same weight to each kernel when combining the basic kernels. The localized multiple kernel learning (LMKL) algorithm uses a gating model to locally select the appropriate weight for each basic kernel. Compared with MKL, LMKL could select suitable weight for the datasets. Experimental results on bioinformatics datasets show that LMKL with the gating model has better accuracy than the model with single kernel [22]. Equation (8) gives a decision function for LMKL [22] as follows:

$$f(x) = \sum_{i=1}^N \alpha_i y_i \sum_{m=1}^P \eta_m(x) K_m(x, x_i) \eta_m(x_i) + b, \quad (8)$$

where N is the number of samples, P is the number of basic kernels, b is the bias, K_m is the m th basic kernel, and $\eta_m(x)$ is the gating model. For input sample x , the gating model chooses feature space m as a function of input sample x . $\eta_m(x)$ can be learned from the sample datasets. The gating model $\eta_m(x)$ is defined as the following equation [22]:

$$\eta_m(x) = \frac{\exp(v_m \cdot x + v_{m0})}{\sum_{k=1}^P \exp(v_k \cdot x + v_{k0})}, \quad (9)$$

where v_m and v_{m0} are the parameters of the gating model and the softmax guarantees nonnegativity. By modifying equation (8), with selection function, we get the following optimization problem as the following equation [23]:

$$\begin{aligned}
& \max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K_\eta(x_i, x_j) \\
& 0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \quad (10) \\
& \text{s.t.} \quad \sum_{i=1}^N \alpha_i y_i = 0,
\end{aligned}$$

where $K_\eta(x_i, x_j)$ is defined as equation (11); here, K_η is positive semidefinite [23].

$$K_\eta(x_i, x_j) = \sum_{m=1}^P \eta_m(x_i) K_m(x_i, x_j) \eta_m(x_j). \quad (11)$$

Derivatives of equation (10) are taken with respect to v_m and v_{m0} , and then we use gradient descent to train the gating model. We just need to fix $\eta_m(x)$ and then solve a canonical multiple kernel SVM dual problem first and then update the parameters of the gating model with gradient descent at each step.

2.7. Kernels Selection and Model Selection. Traditional multiple kernel learning methods only select a few simple basic kernels, such as linear kernel (K_L), polynomial kernel (K_P), and Gaussian (or RBF) kernel (K_G). In our proposed model, we selected three simple basic kernels, linear kernel, polynomial kernel, and Gaussian kernel and a deep synthetic kernel proposed above (K_D) as the localized multiple kernel learning combination. Finally, we got the LMKL-D model as shown in Figure 3. K_D was obtained by the deep multiple kernel learning (DMKL) model. The formulas for the three simple basic kernel functions are shown in the following equation [20]:

$$\begin{aligned}
K_L(x_i, x_j) &= x_i \cdot x_j, \\
K_P(x_i, x_j) &= (\alpha x_i \cdot x_j + \beta)^q, \\
K_G(x_i, x_j) &= \exp\left(-\frac{\|x_i - x_j\|^2}{s^2}\right).
\end{aligned} \quad (12)$$

We used grid search to find the parameters of the simple basic kernels. The parameters with the highest accuracy were adopted. Finally, the parameter of the Gaussian kernel s was set to 1 and the polynomial kernel exponent q was set to 2, while α and β were both 1.

We used multiple kernel learning models to obtain K_D . Multiple kernel learning models often use linear kernels, Gaussian kernels, and polynomial kernels to map input data into feature spaces. Since the DMKL model should try to maximize the upper bound of the final kernel to increase its richness with each layer, we combined the linear kernel, polynomial kernel, and Gaussian kernel into one set of kernels. From [23], the number of layers for the DMKL was set to 3. K_D was trained on train set and tested on test set. We used leave one out validation and the minimum value of span to evaluate K_D . K_D with a minimum span value was

adopted. To find better performance, the penalty parameter C was set in the range of $\{10^{-6}, 10^{-5}, \dots, 10^{-1}, 1, 10\}$ and the learning rate was set in the range of $\{10^{-6}, 10^{-5}, \dots, 10^{-1}, 1, 10\}$. After trained and tested, we got K_D . In the end, we chose four kernels, K_L , K_P , K_G , and the best K_D as the final basic kernels of localized multiple kernel learning.

For model selection, the dataset selection operations were repeated three times, and the average value of the results on test set was taken as the final performance of the model. Thus, for each training and test, the training set had 7,500 samples in total and the test set had 3,100 samples in total. For the DMKL model, we used LIBSVM [36] package to solve the SVM optimization problem. For localized multiple kernel learning, we used SMO to speed up the SVM optimization.

3. Results and Discussion

3.1. Comparison with Other Classification Methods. In order to evaluate the performance of the localized multiple kernel learning using the deep synthetic kernel (LMKL-D) model proposed in this paper, the performances of the LMKL-D model were measured by sensitivity (SE, the proportion of the positive examples correctly classified), specificity (SP, the proportion of the negative examples correctly classified), geometric mean (GM, the square root of the product SE and SP), and accuracy (ACC, the percentage of correctly classified instances). SE, SP, GM, and ACC are defined in equation (13) [20]. We compare the LMKL-D model with triplet-SVM [14], miPred [27], MiPred [37], and a three-layer backpropagation neural network (BPNN). The results are shown in Figure 4 and Table 1. Ultimately, the LMKL-D model obtained an accuracy rate of 98.03% on test set, while the triplet-SVM, miPred, MiPred, and BPNN (3 layers) on the test set obtained accuracy rate of 83.90%, 93.50%, 91.29%, and 95.18%, respectively.

$$\begin{aligned}
SE &= \frac{TP}{TP + FN}, \\
SP &= \frac{TN}{TN + FP}, \\
GM &= \sqrt{SE \times SP}, \\
ACC &= \frac{TP + TN}{TP + TN + FP + FN}.
\end{aligned} \quad (13)$$

As shown in Figure 4 and Table 1, LMKL-D has the best 99.27% SP, best 96.11% GM, and best 98.03% ACC, which means it can better distinguish real pre-microRNAs and pseudo hairpin sequences. Since there are a large number of gene sequences with hairpin structures to be identified, higher specificity can filter the pseudo hairpin sequences. For geometric mean, LMKL-D achieved the highest geometric mean (96.11%) among these methods. That means the LMKL-D model can achieve high performance while maintaining stability. The AUC (area under the curve) 0.9611 (we can see in Figure 5) indicates that the LMKL-D

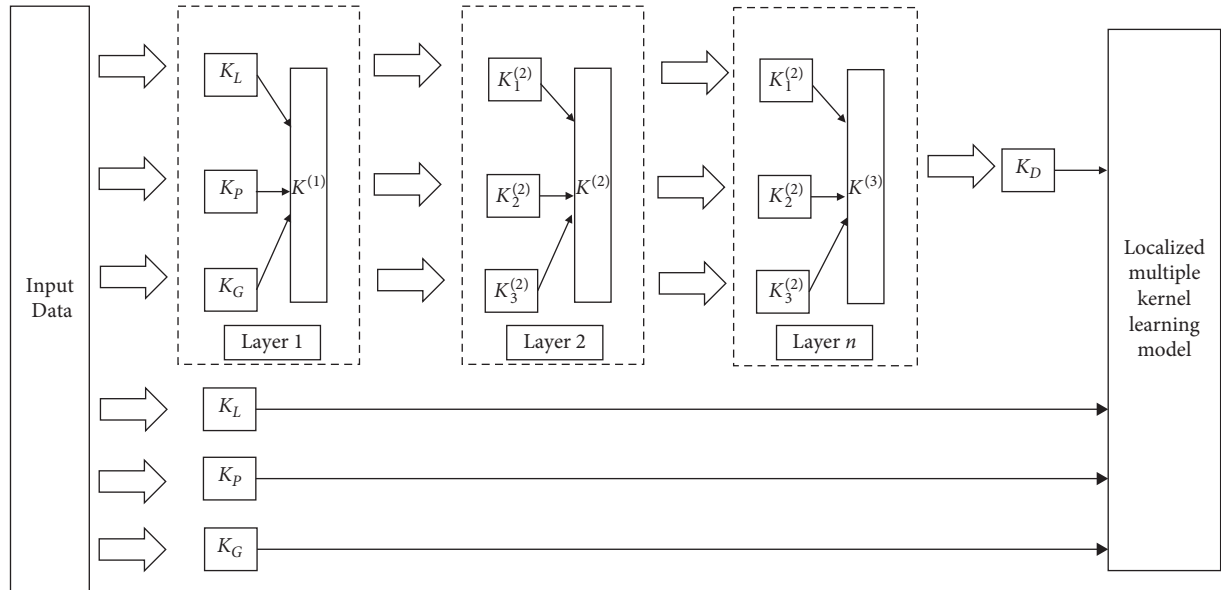


FIGURE 3: LMKL-D model proposed in this paper.

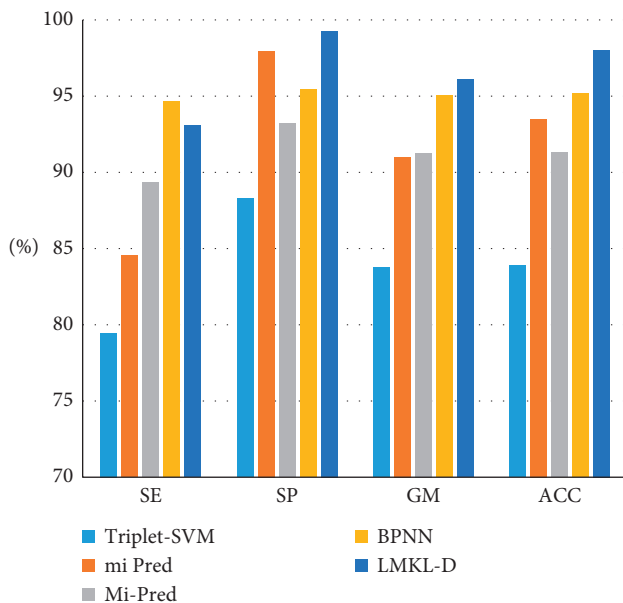


FIGURE 4: Comparison of prediction performance of LMKL-D with other existing methods. The BPNN was three layers here.

TABLE 1: Comparison of prediction performance of LMKL-D with other existing methods.

Methods	SE	SP	GM	ACC
Triplet-SVM (libSVM) [7]	79.47	88.30	83.77	83.90
miPred (libSVM) [19]	84.55	97.97	91.01	93.50
MiPred (random forest) [28]	89.35	93.21	91.26	91.29
BPNN (3 layer)	94.64	95.44	95.04	95.18
LMKL-D	93.06	99.27	96.11	98.03

can predict pre-microRNA accurately. Since the ratio of pre-microRNA sequences to pseudo hairpin sequences is about 1 to 4, the SE of the LMKL-D model might be lower. Next, we

need find new methods to deal with class imbalances. These data prove that our proposed localized multiple kernel learning using the deep synthetic kernel model can increase classification accuracy with low risk of over-fitting and has a more accurate predictive ability and stability to identify the new microRNA precursors in many species.

3.2. Comparison with Localized Multiple Kernel Learning.

In order to better evaluate our LMKL-D model, we also compared LMKL-D with LMKL. For basic kernels, the LMKL-D model used four basic kernels, K_L , K_P , K_G , and K_D . The LMKL model used three basic kernels, K_L , K_P , and K_G . K_L , K_P , and K_G of the two models adopted the same parameters. The penalty parameter C was fixed on 0.035. The results on test set are shown in Figure 6. The performance of the two models on training and test set is shown in Table 2.

From Figure 6 and Table 2, we can see that the LMKL-D model has 91.33% SE, 99.22% SP, 97.64% ACC, and 0.9535 AUC on training set, while the LMKL model obtained 87.47% SE, 99.60% SP, 97.17% ACC, and 0.9352 AUC. On test set, the LMKL-D model has 93.06% SE, 99.27% SP, 98.03% ACC, and 0.9611 AUC, while the LMKL model obtained 88.71% SE, 99.60% SP, 97.42% ACC, and 0.9407 AUC. On both the training set and the test set, the LMKL-D model has 91.83% SE, 99.23% SP, 97.75%, and 0.9574 AUC, while the LMKL model obtained 87.83% SE, 99.60% SP, 97.24% ACC, and 0.9383 AUC. Although LMKL-D acquires a little lower specificity than LMKL, on sensitivity, accuracy, and AUC, LMKL-D is always better than LMKL. On geometric mean, LMKL-D is always higher than LMKL. That means that LMKL-D is more stable than LMLK. We can draw a conclusion from Figure 6 and Table 2 that LMKL-D has better sensitivity, geometric mean, and accuracy than LMKL. For specificity, the two models have similar performance. The results show that in terms of correctly and stably identifying pre-microRNA, the LMKL-D is more

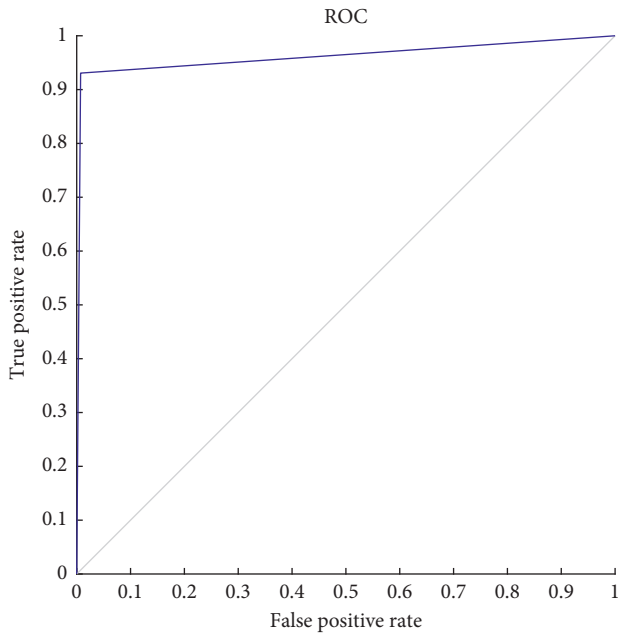


FIGURE 5: ROC curve of LMKL-D; AUC=0.9611.

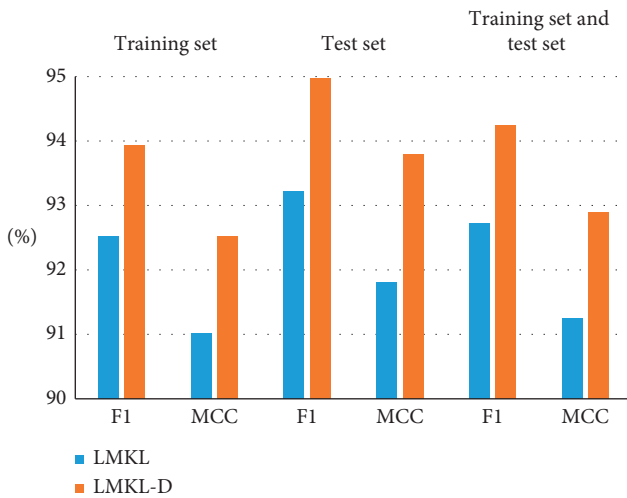


FIGURE 6: Comparison of LMKL-D and LMKL on F1 and MCC.

TABLE 2: Comparison of prediction performance of LMKL-D with LMKL.

Dataset	Methods	SE (%)	SP (%)	GM (%)	ACC (%)	AUC
Training set	LMKL	87.47	99.60	93.34	97.17	0.9352
	LMKL-D	91.33	99.22	95.19	97.64	0.9535
Test set	LMKL	88.71	99.60	94.00	97.42	0.9407
	LMKL-D	93.06	99.27	96.11	98.03	0.9611
Training set And test set	LMKL	87.83	99.60	93.53	97.24	0.9383
	LMKL-D	91.83	99.23	95.46	97.75	0.9574

efficient than LMKL. The experiments also show that the deep model can obtain more features than the LMKL model. The LMKL-D model obtained both deep and shallow features of the samples.

4. Conclusions

In this work, we have proposed a localized multiple kernel learning model with a three-layer deep synthetic kernel in improving the pre-microRNAs prediction accuracy of existing methods. The experiments show that our proposed model yielded comparable better predictive performances and is more stable than existing classifiers for identifying known pre-microRNAs. After being trained on hairpin sequences train set, the LMKL-D methods obtain 93.06% sensitivity, 99.27% specificity, 96.11% geometric mean, and 98.03% accuracy on test set. By applying deep architecture to localized multiple kernel learning, we found that the LMKL-D model is both useful and reliable, as demonstrated in the results above. The LMKL-D model was examined by comparing with the triplet-SVM, miPred, MiPred, and a three-layer backpropagation neural network. We also compare the LMKL-D model and LMKL model. Our results show a more efficient model compared with the multiple kernel learning model with simple basic kernels. The three-layer deep synthetic kernel can indeed increase the richness of kernels and represent deep features. On the other hand, the LMKL-D model could use both shallow and deep features. The number of pseudo hairpin sequences in nature is much larger than known pre-microRNAs. There are always more negative samples than positive samples. With the development of bioinformatics, it is still a challenging work to solve the problem of sample imbalance and explore more classification methods.

Data Availability

The known pre-microRNA sequences data can be downloaded from the miRBase website at <http://www.mirbase.org>; the UCSC refGene annotation list and the human RefSeq gene were available through <https://www.ncbi.nlm.nih.gov/refseq/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Funds (grant nos. 62072391, 61872419, and 61573166), Joint Funds of ShanDong Natural Science Funds (grant no. ZR2019LZH003), Taishan Scholars Program of Shandong Province, China (grant no. tsqn201812077), Shandong Provincial Natural Science Foundation (grant no. ZR2018LF005), University Innovation Team Project of Jinan (grant no. 2019GXRCO15), and Key Science and Technology

Innovation Project of Shandong Province (grant no. 2019JZZY010324).

References

- [1] A. E. Erson-Bensan, "Introduction to microRNAs in biological systems," *Methods in Molecular Biology*, vol. 1107, pp. 1–14, 2014.
- [2] D. Zhengren, Y. Xiuhong, W. Yuqiang, and W. yanjin, "Research progress of miRNA related to hyperlipidemia induced coronary arterial endothelial injury," *Chinese Journal of Arteriosclerosis*, vol. 28, no. 8, pp. 721–727, 2020.
- [3] J. X. Wang Yuhan, "Research progress in biological function of microRNA-129," *Military Medical Sciences*, vol. 42, no. 5, pp. 384–387, 2018.
- [4] L. Wei, M. Liao, Y. Gao, R. Ji, Z. He, and Q. Zou, "Improved and promising identification of human microRNAs by incorporating a high-quality negative set," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 11, no. 1, pp. 192–201, 2013.
- [5] L. Jiang, J. Zhang, P. Xuan, and Q. Zou, "BP neural network could help improve pre-miRNA identification in various species," *BioMed Research International*, vol. 2016, Article ID 9565689, 2016.
- [6] X. Zeng, L. Liu, L. Lü, and Q. Zoü, "Prediction of potential disease-associated microRNAs using structural perturbation method," *Bioinformatics*, vol. 34, no. 14, pp. 2425–2432, 2018.
- [7] W. Bao, D.-S. Huang, and Y.-H. Chen, "Msit: malonylation sites identification tree," *Current Bioinformatics*, vol. 15, no. 1, pp. 59–67, 2020.
- [8] W. Bao, B. Yang, D.-S. Huang et al., "IMKPse: identification of protein malonylation sites by the key features into general PseAAC," *IEEE Access*, vol. 7, pp. 54073–54083, 2019.
- [9] R. C. Lee, R. L. Feinbaum, and V. Ambros, "The *c. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*," *Cell*, vol. 75, no. 5, pp. 843–854, 1993.
- [10] F. Huang, X. Yue, Z. Xiong, Z. Yu, S. Liu, and W. Zhang, "Tensor decomposition with relational constraints for predicting multiple types of microRNA-disease associations," *Briefings in Bioinformatics*, 2019.
- [11] W. Zhang, Z. Li, W. Guo, W. Yang, and F. Huang, "A fast linear neighborhood similarity-based network link inference method to predict microRNA-disease associations," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 2, 2019.
- [12] Y. Gong, Y. Niu, W. Zhang, and X. Li, "A network embedding-based multiple information integration method for the miRNA-disease association prediction," *BMC Bioinformatics*, vol. 20, no. 1, pp. 1–13, 2019.
- [13] Y. Deng, X. Xu, Y. Qiu, J. Xia, W. Zhang, and S. Liu, "A multimodal deep learning framework for predicting drug-drug interaction events," *Bioinformatics*, vol. 36, no. 15, pp. 4316–4322, 2020.
- [14] C. Xue, F. Li, T. He, G.-P. Liu, Y. Li, and X. Zhang, "Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine," *BMC Bioinformatics*, vol. 6, no. 1, p. 310, 2005.
- [15] S. G. Das, H. J. Chakraborty, and A. Datta, "PremipreD: precursor miRNA prediction by support vector machine approach," *Trends in Bioinformatics*, vol. 11, no. 1, pp. 17–24, 2018.
- [16] S.-Y. Han, J. Zhou, Y.-H. Chen, Y.-F. Zhang, G.-Y. Tang, and L. Wang, "Active fault-tolerant control for discrete vehicle active suspension via reduced-order observer," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.
- [17] W. Bao, B. Yang, D. Li, Z. Li, Y. Zhou, and R. Bao, "Cmsenn: computational modification sites with ensemble neural network," *Chemometrics and Intelligent Laboratory Systems*, vol. 185, pp. 65–72, 2019.
- [18] D. Zheng, J. Wang, and Y. Zhao, "Non-flat function estimation with a multi-scale support vector regression," *Neurocomputing*, vol. 70, no. 1–3, pp. 420–429, 2006.
- [19] R. J. Peace, M. S. Hassani, and J. R. Green, "Miple: NGS-based prediction of miRNA using integrated evidence," *Scientific Reports*, vol. 9, no. 1, p. 1548, 2019.
- [20] H. Shi, W. Wang, P. Wu, and D. Wang, "Support vector machine based on localized multiple kernel learning in pre-microRNA classification," in *Proceedings of the 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pp. 1–5, Istanbul, Turkey, June 2020.
- [21] M. E. Rahman, R. Islam, S. Islam, S. I. Mondal, and M. R. Amin, "MiRANN: a reliable approach for improved classification of precursor microRNA using artificial neural network model," *Genomics*, vol. 99, no. 4, pp. 189–194, 2012.
- [22] M. Gönen and E. Alpaydin, "Localized multiple kernel learning," in *Proceedings of the 25th International Conference on Machine Learning*, pp. 352–359, Helsinki, Finland, July 2008.
- [23] E. V. Strobl and S. Visweswaran, "Deep multiple kernel learning," in *Proceedings of the 2013 12th International Conference on Machine Learning and Applications*, vol. 1, pp. 414–417, Miami, FL, USA, December 2013.
- [24] A. Kozomara and S. Griffithsjones, "MiRBase: annotating high confidence microRNAs using deep sequencing data," *Nucleic Acids Research*, vol. 42, no. 1, pp. 68–73, 2014.
- [25] D. Karolchik, R. M. Kuhn, R. Baertsch et al., "The UCSC genome browser database: 2008 update," *Nucleic Acids Research*, vol. 36, pp. 773–779, 2007.
- [26] D. Maglott, J. Ostell, K. D. Pruitt, and T. Tatusova, "Entrez gene: gene-centered information at NCBI," *Nucleic Acids Research*, vol. 33, pp. 26–31, 2004.
- [27] K. L. S. Ng and S. K. Mishra, "De novo SVM classification of precursor microRNAs from genomic pseudo hairpins using global and intrinsic folding measures," *Bioinformatics*, vol. 23, no. 11, pp. 1321–1330, 2007.
- [28] W. Bao, C.-A. Yuan, Y. Zhang et al., "Mutli-features prediction of protein translational modification sites," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 5, pp. 1453–1460, 2017.
- [29] R. Lorenz, S. H. Bernhart, C. H. Z. Siederdisen et al., "ViennaRNA package 2.0," *Algorithms for Molecular Biology*, vol. 6, no. 1, p. 26, 2011.
- [30] M. Gönen and E. Alpaydin, "Multiple kernel learning algorithms," *The Journal of Machine Learning Research*, vol. 12, pp. 2211–2268, 2011.
- [31] H. Wang and S. Fuchun, "On multiple kernel learning methods," *Acta Automatica Sinica*, vol. 36, no. 8, pp. 1038–1050, 2010.
- [32] Y. Liu, S. Liao, and Y. Hou, "Learning kernels with upper bounds of leave-one-out error," in *Proceedings of the 20th ACM International Conference on Information and knowledge management*, pp. 2205–2208, Glasgow, Scotland, October 2011.
- [33] W.-J. Lee, S. Verzakov, and R. P. Duin, "Kernel combination versus classifier combination," in *Proceedings of the 7th*

International Workshop on Multiple Classifier Systems, pp. 22–31, Prague, Czech Republic, May 2007.

- [34] W. Bao, D. Wang, and Y. Chen, “Classification of protein structure classes on flexible neutral tree,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 14, no. 5, pp. 1122–1133, 2016.
- [35] O. Chapelle and V. Vapnik, “Model selection for support vector machines,” *Advances in Neural Information Processing Systems*, vol. 12, pp. 230–236, 1999.
- [36] C.-C. Chang and C.-J. Lin, “Libsvm: a library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.
- [37] P. Jiang, H. Wu, W. Wang, W. Ma, X. Sun, and Z. Lu, “MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features,” *Nucleic Acids Research*, vol. 35, pp. 339–344, 2007.

Research Article

A Novel Transfer Enhanced α -Expansion Move Learning Model for EEG Signals

Jiangwei Cai , Lu Zhao , and Anqi Bi 

Department of Computer Science and Engineering, Changshu Institute of Technology, Changshu, Jiangsu, China

Correspondence should be addressed to Anqi Bi; anqi_b@cslg.edu.cn

Received 16 March 2021; Revised 9 April 2021; Accepted 16 April 2021; Published 24 April 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Jiangwei Cai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we focus on recognizing epileptic seizure from scant EEG signals and propose a novel transfer enhanced α -expansion move (TrEEM) learning model. This framework implants transfer learning into the exemplar-based clustering model to improve the utilization rate of EEG signals. Starting from Bayesian probability theory, by leveraging Kullback-Leibler distance, we measure the similarity relationship between source and target data. Furthermore, we embed this relationship into the calculation of similarity matrix involved in the exemplar-based clustering model. Then we sum up a new objective function and study this new TrEEM scheme earnestly. We optimize the proposed TrEEM model by borrowing the mechanism utilized in EEM. In contrast to other machine learning models, experiments based on synthetic and real-world EEG datasets show that the performance of the proposed TrEEM is very promising.

1. Introduction

Epilepsy is a kind of chronic disease, which is caused by the sudden abnormal discharge of brain neurons, resulting in transient brain dysfunction. Usually patients themselves have no obvious impression of the epileptic seizure process. For this reason, doctors can only diagnose the patient's condition according to the patient's family members or other personnel present during the epileptic seizure in the past. However, the accuracy of this manual diagnosis method is low. The pathogenesis of epilepsy is mainly manifested by abnormal neural discharge and abnormal brain waves. Although medical imaging, such as Computed Tomography (CT), magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), Single-Photon Emission Computed Tomography (SPECT), Positron Emission Computed Tomography (PET), has made great progress over the years, and the major diagnostic method of epilepsy is based on electroencephalogram (EEG). More specifically, PET and fMRI cannot be used as common technical means because of their technical requirements and costs. In addition to the high cost, MRI cannot judge the nonstructural lesions as well. Invasive cortical electroencephalogram (ECoG) requires

craniotomy and implantation of electrodes, which has a high risk; and noninvasive EEG and MEG can provide functional and structural detection. Taking all these into account, EEG has been widely concerned in more and more theoretical researches and clinical practice because of its low cost, convenient signal acquisition, and noninvasiveness.

The research on diagnosis of epilepsy through EEG signals has been a hot topic in related fields, compared with manual diagnostic method, and machine learning methods are less time-consuming and more accurate [1–8]. Numerous machine learning models have been used to recognize epileptic EEG signals, such as support vector learning [9, 10], fuzzy system [1, 3], naïve Bayes [11], and exemplar-based clustering model [2, 12, 13]. The traditional machine learning process is usually divided into the three following steps, as shown in Figure 1: (1) EEG signal preprocessing improves signal to noise ratio and provides high-quality input signal for spike detection. (2) According to the characteristics of spikes, artificial design features can reduce the signal dimension and highlight the difference between spikes and background signals. (3) According to the obtained features, spike signals are detected by the machine learning mechanism involved.

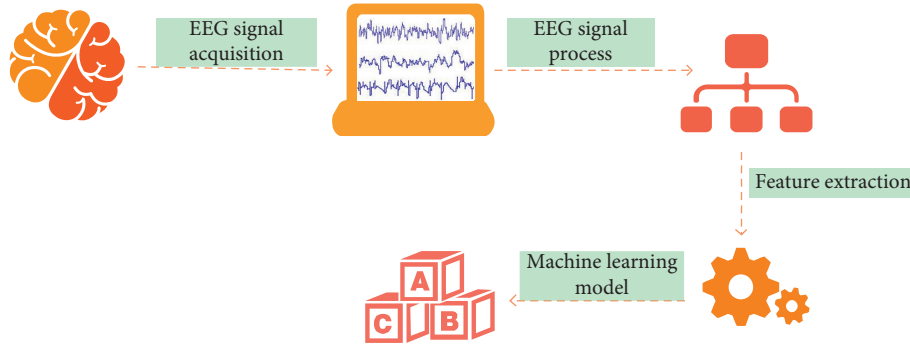


FIGURE 1: General steps of machine learning model processing EEG signals.

In summary, one of the significant issues in the field of processing EEG signals by machine learning technique is the insufficient training data. We briefly introduce some mechanisms for epileptic diagnosis through EEG signals here. Jiang [1] integrates transductive transfer learning, semisupervised learning, and Takagi-Sugeno-Kang (TSK) fuzzy system to take full advantage of the scant training data. Zhu [5] proposed dic-mv-fcm, which automatically evaluates the importance and weights of each view and then performs weighted multiview fuzzy clustering based on FCM framework to achieve accurate fuzzy partition. Bi [2] proposed a novel model called FEEM for incomplete EEG signal, which first compresses the potential exemplar list and thus reduces the scale of pairwise similarity matrix. However, to make better use of training data, we still need to do a lot of work, and we focus on this issue in this paper as well. Specifically, this paper aims at recognizing epileptic seizure from scant EEG signals.

Transfer learning is believed to be an effective strategy to solve problems caused by insufficient training data [1, 5, 13, 14]. Assume that there are two datasets from similar source: one has plenty of features and details and is easy to be learned, while the other one lacks details and is hard to be learned. Transfer learning offers an idea of leveraging the description of former data to study the latter data. The sufficient well-described data is called source data, while the insufficient rough data is named as target data. Accordingly, transfer learning utilizes source data to improve the learning result of target data. Under this framework, effectively measuring the relationship between source data and target data is an important part and has a great influence on the efficiency of relevant study model. Thus, starting from Bayesian probability theory, this paper first extends the concept of similarity matrix in the exemplar-based clustering model; and this strategy also broadens the application range of the algorithm to transfer learning scenario. By leveraging Kullback–Leibler distance, we propose a new transfer enhanced α -expansion move learning model called TrEEM. The detailed contributions of this paper are listed as follows:

- (i) According to the transfer learning theory [1, 5, 13, 14], considering the similarity between source and target data, the proposed model TrEEM should keep the target data close enough to the

source data. Theoretically supported by the information theory, based on the Bayesian probability framework, TrEEM utilizes KL distance to measure the similarity between source data and target data and minimizes this KL distance in the optimization process.

- (ii) In the scenario of recognizing epileptic seizure, we aim at diagnosing the actual patient. As TrEEM is built on graph theory and pairwise similarity matrix and is an exemplar-based clustering model, this model selects exemplar from actual data. This advantage fits the requirements in the relevant scenario here.
- (iii) TrEEM embeds KL distance between target data and source data into the calculation of similarity matrix. Thus, the optimization mechanism utilized in EEM can be directly used to solve the new target function of TrEEM. In detail, we leverage α -expansion move optimization algorithm which performs better than LBP [15, 16] algorithm.

The paper is organized as follows. The related works are discussed in Section 2. We illustrate the target function and optimization mechanism of the proposed TrEEM in Section 3. The simulation experimental results and analysis are shown in Section 4. We make a conclusion in Section 5.

2. Related Works

Many researchers are committed to using machine learning technology to classify EEG signals, including SVM, fuzzy system, naïve Bayes, and exemplar-based clustering model. In this section, we illustrate two popular learning frameworks, namely, Enhanced α -Expansion Move (EEM) and TSK fuzzy system. EEM is a widely used exemplar-based learning model, and TSK fuzzy system is a typical fuzzy-rule-based clustering model.

2.1. Enhanced α -Expansion Move. Consider a dataset $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{N \times D}$; N is the total number of D -dimensional data points. E is the output, whereas the element $E(\mathbf{x}_i)$ refers to the exemplar for each \mathbf{x}_i .

The target function of a typical exemplar-based clustering model is defined as follows [12, 15]:

$$\max_E \sum_{p=1}^N s(\mathbf{x}_p, \mathbf{x}_{E(p)}) - \sum_{p=1}^N \sum_{q>p}^N \theta_{p,q}(E(p), E(q)), \quad (1)$$

where S is the similarity matrix of the dataset, and the elements are defined as $S(i, j) = -\|\mathbf{x}_i - \mathbf{x}_j\|^2$; $\theta_{p,q}(E(p), E(q))$ is shown as follows:

$$\theta_{p,q}(E(p), E(q)) = \begin{cases} M, & E(p) = q, \quad E(q) \neq p, \\ \text{or } E(q) = p, & E(p) \neq q, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

In [15], the authors regard the above target function as the energy function of Markov random field (MRF) and verifies the below theorem.

Theorem 1. When, for $\forall p, q, E(p), E(q), \alpha \in \{1, 2, \dots, N\}$, equation (3) is verified, the graph-theory based framework can be used to optimize the target function of the exemplar-based clustering model as shown in equation (1).

$$\theta_{p,q}(E(p), E(q)) + \theta_{p,q}(\alpha, \alpha) \leq \theta_{p,q}(E(p), \alpha) + \theta_{p,q}(\alpha, E(q)). \quad (3)$$

Enhanced α -Expansion Move (EEM) framework optimizes the above target function by an improved algorithm [15]. In more detail, theoretically supported by Theorem 1 and graph-cuts [16] algorithm, EEM expands the active region of candidate exemplar from a single data to the whole dataset. EEM defines the second optimal candidate exemplar $S(i)$ for \mathbf{x}_i as below, which is selected from the whole dataset as mentioned above.

$$S(i) = \arg \max_{s \in (E/l)} s(\mathbf{x}_i, \mathbf{x}_s), \quad \forall \mathbf{x}_i \in \mathbf{X}_l, \quad (4)$$

where $\mathbf{X}_l = \{\mathbf{x}_i | E(i) = l\}$ is the dataset among which the exemplar is l , and $s \in (E/l)$ represent other exemplars in E expect for l . In this way, the optimization mechanism behaves more rapidly and effectively.

EEM algorithm is one of the most popular exemplar-based clustering models, and it performs effectively and steadily in numerous simulation experiments involved [2, 12–15]. Scientists have applied this model for data stream, constrained supervised learning, and EEG signal processing.

2.2. TSK Fuzzy System. TSK fuzzy system is a rule-based system and it is widely used as a typical fuzzy system model for both classification and clustering. Generally, the k th TSK fuzzy rule for K fuzzy rules can be described as R^k .

$$\begin{aligned} \text{IF } \mathbf{x}_1 \text{ is } A_1^k \wedge \mathbf{x}_2 \text{ is } A_2^k \wedge \dots \wedge \mathbf{x}_N \text{ is } A_N^k, \\ \text{THEN } f^k(\mathbf{x}) = p_0^k + p_1^k \mathbf{x}_1 + \dots + p_N^k \mathbf{x}_N, \end{aligned} \quad (5)$$

where A_i^k is a fuzzy set subscribed by the input x_i for the k th fuzzy rule and \wedge is a fuzzy conjunction operator. Each rule is premised on the input data $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{N \times D}$ which is mapped to a singleton $f^k(\mathbf{x})$. Thus, the output of the TSK fuzzy system is defined as

$$y^0 = \sum_{k=1}^K \frac{\mu^k(\mathbf{x}) f^k(\mathbf{x})}{\sum_{m=1}^K \mu^m(\mathbf{x})} = \sum_{k=1}^K \tilde{\mu}^k(\mathbf{x}) f^k(\mathbf{x}), \quad (6)$$

where

$$\begin{aligned} \mu^k(\mathbf{X}) &= \prod_{i=1}^N \mu_{A_i^k}(x_i), \\ \tilde{\mu}^k(\mathbf{X}) &= \frac{\mu^k(\mathbf{X})}{\sum_{m=1}^K \mu^m(\mathbf{X})}, \end{aligned} \quad (7)$$

where $\mu_{A_i^k}(x_i)$ is the membership grade that can be obtained using Gaussian membership function, and the other involved parameters also could be estimated using clustering techniques and other partition methods [1, 3–5].

Accordingly, based on the relevant theory of TSK fuzzy system, the target model above in equation (6) converts to a parameter learning process of the corresponding linear regression model. In line with recent achievements, TSK fuzzy model has strong interpretability and robustness. For this reason, this TSK fuzzy model is widely used among numerous intelligent medical diagnosis systems, including recognizing epileptic seizure from EEG signals.

In this section, we briefly introduce two popular machine learning clustering frameworks used in the recognition of EEG signals, namely, EEM and TSK fuzzy system. The detailed descriptions are shown in Table 1. Considering the scenario of diagnosing epileptic patients from some healthy patients based on their EEG signals, we focus on EEM clustering model in the rest of this paper.

3. Transfer Enhanced α -Expansion Move Learning Model

In this section, we first analyze the theoretical basis of TrEEM from Bayesian probabilistic framework. Second, we induce the novel algorithm TrEEM in detail. Then, considering the optimization algorithm utilized in EEM algorithm, we optimize target function as well. Generally, the structure of this novel model is shown in Figure 2. See Figure 2; on the basis of source-data-based exemplar set, starting from Bayesian probability framework, TrEEM first imbeds the distance between source data and target data in the calculation of similarity matrix. This distance is measured by Kullback-Leibler distance. Then we induce the novel target function for TrEEM. Finally, TrEEM directly calls the optimization algorithm in EEM to solve this model and obtain the target-data-based exemplar set.

Besides, we list the frequently used notations in Table 2.

3.1. Theoretical Preliminary of TrEEM Scheme. As mentioned before, transfer learning considers two datasets from similar source, namely, source data and target data; and the relationship between source data and target data is considered as a significant factor in this model (see Table 2); in the following part, we define the sufficient well-described source data as $\tilde{\mathbf{X}}$. After study, we obtain the source-data-based exemplar set denoted as L_s in the above table. Then the insufficient target

TABLE 1: Descriptions of two popular machine learning algorithms used in recognition of EEG signals.

Algorithms	Theoretical basis	Descriptions	Optimization frameworks
EEM	Graph theory	Select exemplar from actual data, do not need to preset the cluster number	Enhanced graph-cuts optimization algorithm, expand the candidate region
TSK	Fuzzy system	Rule-based learning model, strong interpretability and robustness	Parameter learning process of corresponding linear regression model

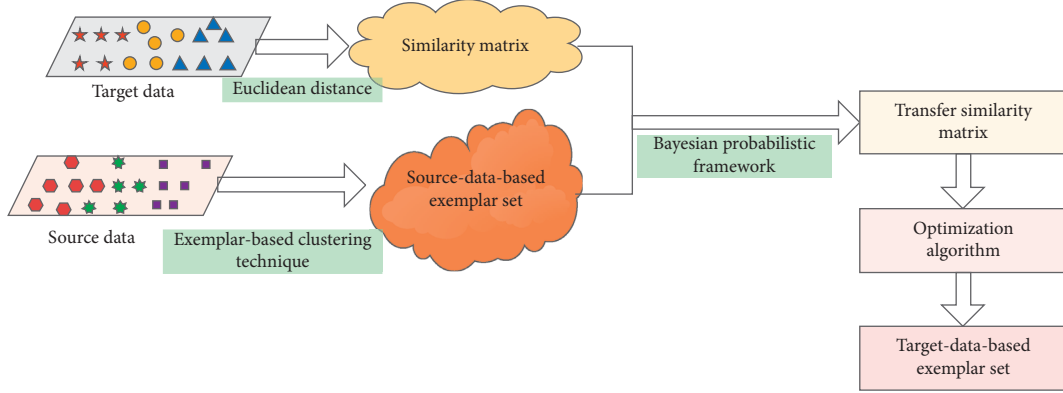


FIGURE 2: Structure of TrEEM algorithm.

TABLE 2: Involved notations and descriptions.

Notations	Descriptions
$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{N \times D}$	Target data
$\tilde{\mathbf{X}} = \{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_{N_s}\} \in \mathbb{R}^{N_s \times D}$	Source data
$\mathbf{S} = (s(\mathbf{x}_i, \mathbf{x}_j))$	Pairwise similarity matrix
L_s	Source-data-based exemplar set
$L_s(i)$	Source exemplar for target sample x_i
E	Target-data-based exemplar set
$E(i)$	Exemplar for target sample x_i

data is defined as \mathbf{X} above. Moreover, probabilistic framework contributes to measuring this relationship as well. Therefore, supported by Gaussian probability hypothesis and exemplar-based cluster mechanism, we built the pairwise probabilistic relationship of target data by leveraging the corresponding similarity as follows:

$$p(x_i, x_{E(i)}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s(x_i, x_{E(i)})}{2\sigma^2}\right), \quad (8)$$

where $s(x_i, x_{E(i)})$ is the similarity between x_i and its current exemplar $x_{E(i)}$ and parameter σ is a standard deviation from Gaussian probability hypothesis.

As to the exemplar set, we should exclude the situation when an exemplar appoints other exemplars among current exemplar set except for itself as its own exemplar. Consequently, Bayesian posterior probability of an exemplar set is defined as follows:

$$p(E) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\sum_{m=1}^N \sum_{n=1}^N \frac{\theta_{m,n}(E(m), E(n))}{2\sigma^2}\right), \quad (9)$$

and $\theta_{m,n}(E(m), E(n))$ here is the same as the definition shown in equation (2).

Accordingly, under Bayesian probabilistic framework and the discussion of EEM algorithm in Section 2, the objective function in equation (1) is equal to the following function:

$$\begin{aligned} \max_E Q(E) &= \prod_{i=1}^N \ln p(x_i, x_{E(i)}) \ln p(E) \\ &= \sum_{i=1}^N \ln p(x_i, x_{E(i)}) + \ln p(E). \end{aligned} \quad (10)$$

In conclusion, equation (10) defines another form of the target function of EEM by introducing Bayesian probabilistic framework and Gaussian probability hypothesis. Starting from this target function, we would be able to design TrEEM for recognition of epileptic EEG signals in the next subsections.

3.2. Design of TrEEM Scheme. According to information theory, the Kullback-Leibler distance (KL distance) is a natural distance between two real probability distributions and it has been widely applied to solve numerous issues [17–19]. The definition of KL distance is shown below.

Definition 1. Consider two probability distributions as P and Q ; the KL distance from P to Q is as follows:

$$D_{KL}(P||Q) = \sum_{x \in X} P(x) \ln \frac{P(x)}{Q(x)}, \quad (11)$$

where $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$ is the input data.

What is worth mentioning is the fact that KL distance is an asymmetric measurement, namely, $D_{KL}(P\|Q) \neq D_{KL}(Q\|P)$, according to Definition 1.

Furthermore, given L_s as a possible exemplar set, $L_s(i)$ is the best exemplar for x_i among current exemplar set L_s . As discussed above, we also define $L_s(x_i)$ under Bayesian probabilistic framework as follows:

$$L_s(i) = \arg \max_{l \in L_s} p(x_i, \tilde{x}_l), \quad (12)$$

where $p(x_i, \tilde{x}_l)$ is obtained from equation (8). In transfer learning, actually two datasets are involved, that is, source data and target data. In equation (12), note that the first x_i is from the target data, and the second \tilde{x}_l is from the possible exemplar set, namely, from the source data. Thus, see Table 2; to make the distinction clear, the symbol \tilde{x}_l represents the source data, while x_i stands for the target data in the rest of this paper.

Although the target data is not exactly same as source data, according to those theoretical analyses of transfer learning, the source-data-based learning model and results should contribute to the learning of new target data as well [3, 4, 20–22]. Otherwise, it will become negative transfer

learning, which is not under discussion in this paper. Accordingly, we believe the target-data-based exemplar that is set to be evaluated is supposed to be similar to the source-data-based exemplar set. In this section, we measure the difference between target-data-based exemplar set and source-data-based exemplar set by the aforementioned KL distance with the definition shown in Definition 1. To be specific, in the process of designing the TrEEM learning model, we minimize the difference of target exemplar E and source exemplar set L_s by controlling the KL distance between them. The structure of TrEEM is shown in Figure 2 in detail. In view of this goal, on the basis of the probabilistic target function in equation (10) of EEM, we build the novel target function for the proposed TrEEM model as follows:

$$\max Q(E) = \frac{\prod_{i=1}^N \ln p(x_i, x_{E(i)}) \ln p(E)}{\lambda D_{KL}(L_s, E)}, \quad (13)$$

where E is the target-data-based exemplar set to be obtained and L_s represents the source-data-based exemplar set, as shown in Table 2. λ is the regularization parameter. In terms of maximum a priori (MAP) principle and combining Definition 1 and equation (12), (13) becomes

$$\max Q(E) = \sum_{i=1}^N \ln p(x_i, x_{E(i)}) + \ln p(E) - \lambda \sum_{i=1}^N p(x_i, \tilde{x}_{L_s(i)}) [\ln p(x_i, \tilde{x}_{L_s(i)}) - \ln p(x_i, x_{E(i)})]. \quad (14)$$

Observing equation (14), we can find that the values of the second and third terms belong to the same magnitudes; hence, the value of λ will not be large and the specific determination strategy will be discussed in Section 4.

Introducing the definitions of $p(x_i, x_{E(i)})$, $p(x_i, \tilde{x}_{L_s(i)})$, and $p(E)$ in equations (8) and (9) and discarding the constant terms, equation (14) can be simplified into the following equation:

$$\max Q(E) = \ln p(E) + \sum_{i=1}^N [1 + \lambda p(x_i, \tilde{x}_{L_s(i)})] \ln p(x_i, x_{E(i)}). \quad (15)$$

Comparing equations (15) and (10), we conclude that they are similar in structure. According to Theorem 1 in Section 2, TrEEM's target function also can be solved by graph-cuts mechanism. Consequently, we will discuss the optimization mechanism step by step in the next subsection.

3.3. Optimization Mechanism of the TrEEM Scheme. As mentioned before, the novel target function in equation (15) is similar to that of EEM algorithm under Bayesian probabilistic framework, so the optimization mechanism utilized in the EEM algorithm is supposed to be helpful in solving the novel target function. However, we need to deal with the difference between these two models firstly.

In detail, we redefine the similarity relationship of target data by imbedding source-data-based exemplar set L_s .

Specifically, we single out the suitable exemplar from L_s for target sample x_i by equation (12) and build the new pairwise transfer similarity matrix $\mathbf{S}_t = (s_t(x_i, x_j))$ according to the new measurement in the following equation:

$$s_t(x_i, x_j) = -\left[1 + \lambda p(x_i, \tilde{x}_{L_s(x_i)})\right] d(x_i, x_j), \quad (16)$$

where $d(x_i, x_j) = \|x_i - x_j\|^2$ is the Euclidean distance between samples x_i and x_j , λ is the regularization parameter, $\tilde{x}_{L_s(x_i)}$ refers to the exemplar singled out from source data. By introducing this new definition of similarity relationship, the target function equation (15) of TrEEM is equal to equation (10) in structure. Meanwhile, the constraint condition in Theorem 1 is true for TrEEM model as well. Therefore, the optimization mechanism of EEM algorithm is also suitable for the proposed TrEEM model; and the novel model TrEEM is described in detail in Algorithm 1.

EEM utilizes α -expansion move to optimize its learning model. As discussed above, the mechanism is also suitable for the proposed TrEEM model. We analyze this Enhanced α -Expansion Move optimization mechanism step by step here. Firstly, as the target functions shown in both equations (15) and (10) also can be defined as the energy function of MRF, we consider this optimization process as an energy reduction process of the MRF. In general, we start from the change values of energy to decide whether to accept new exemplar for a sample. Secondly, the improved optimization mechanism is designed to broaden the effective field when changing the sample's exemplar. That is to say, assume that a

Input: Target dataset $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{N \times D}$, source data $\tilde{\mathbf{X}} = \{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_{N_s}\} \in \mathbb{R}^{N_s \times D}$, source-data-based exemplar set L_s , self similarity $d(\mathbf{x}_i, \mathbf{x}_j)$, regularization factor λ, σ .

Output: Valid target-data-based exemplar set $E(N)$.

- (1) **for** $\mathbf{x}_i \in \mathbf{X}$ **do**
- (2) single out the nearest exemplar $L_s(i)$ for \mathbf{x}_i from source-data-based exemplar set L_s based on equation (12).
- (3) compute probabilistic Euclidean similarity $p(\mathbf{x}_i, \tilde{\mathbf{x}}_{L_s(i)})$ between \mathbf{x}_i and $L_s(i)$.
- (4) **end**
- (5) **for** $\mathbf{x}_i \in \mathbf{X}$ **do**
- (6) calculate transfer similarity matrix $\mathbf{S}_t = (s_t(\mathbf{x}_i, \mathbf{x}_j))$ by new probabilistic similarity $p(\mathbf{x}_i, \tilde{\mathbf{x}}_{L_s(i)})$ according to equation (16).
- (7) **end**
- (8) call the optimization process of EEM as shown in Algorithm 2.

ALGORITHM 1: Transfer Enhanced α -Expansion Move learning model.

Input: similarity matrix \mathbf{S} , dataset $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{N \times D}$, maximum number of iteration M

Output: valid exemplar set $E(N)$.

- (1) Randomly generate expansion order o .
- (2) Let $t = 1$;
- (3) **for** $e \in o$ **do**
- (4) **if** $e \in E$ **then**
- (5) compute R_{in}, R_{in}^l by equations (18), (19)
- (6) **if** $R_{in}^l > 0$ **then**
- (7) for $\forall \mathbf{x}_i \in \mathbf{X}_e$, set $E(\mathbf{x}_i) = A(i)$
- (8) **end**
- (9) **else**
- (10) compute $R_{out}, R_{out}^e, R_{out}^l$ by equations (20), (21), (22)
- (11) **if** $R_{out}^l > R_{out}^e$ **then**
- (12) for $\forall \mathbf{x}_i \in \mathbf{X}_{e,l}^e$, set $E(\mathbf{x}_i) = l$
- (13) **else**
- (14) for $\forall \mathbf{x}_i \in \mathbf{X}_e$, set $E(\mathbf{x}_i) = A(i)$
- (15) **end**
- (16) **if** $R_{out} > 0$ **then**
- (17) Accept the new exemplar l
- (18) **end**
- (19) **end**
- (20) $t = t + 1$
- (21) **end**
- (22) Until convergence

ALGORITHM 2: Optimization model of Enhanced α -Expansion Move algorithm.

sample's current exemplar is abandoned; it will search all the rest exemplars for a new exemplar. This new alternative exemplar is defined as follows:

$$A(i) = \arg \max_{a \in (E/l)} s(\mathbf{x}_i, \mathbf{x}_a), \quad \forall \mathbf{x}_i \in \mathbf{X}_l, \quad (17)$$

where l is the original exemplar for \mathbf{x}_i , E is current exemplar set, and $a \in (E/l)$ is the obtained alternative exemplar. By introducing this alternative exemplar $A(i)$ for \mathbf{x}_i , we enhance the optimization efficiency.

Note that TrEEM model redefines the similarity matrix as equation (16). So, the following discussion is based on the similarity matrix $\mathbf{S}_t = (s_t(\mathbf{x}_i, \mathbf{x}_j))$. Apparently, the optimization mechanism would consider two cases; namely, \mathbf{x}_i is among current exemplar set or is not among

current exemplar set. We analyze these two cases step by step in the next subsections.

3.3.1. Case I. \mathbf{x}_i is a current exemplar.

Obviously, in the process of optimization, this current exemplar \mathbf{x}_i may be abandoned. As previously analyzed, whether to keep \mathbf{x}_i in the ultimate exemplar set is decided by the reduction values of energy function calculated by the target function in equation (15).

Specifically, if \mathbf{x}_i is accepted as an exemplar, the energy of the model remains unchanged, and the reduction value is equal to 0. Otherwise, if \mathbf{x}_i is not accepted, all samples whose exemplars are l would redetermine their exemplars; these samples are defined as $\mathbf{X}_l = \{\mathbf{x}_i | E(i) = l\}$. Theoretically supported by the related analysis in [2, 12, 14, 15], new exemplar for $\mathbf{x}_i \in \mathbf{X}_l$

would be $A(i)$ as shown in equation (17). Thus, the energy reduction R_{in}^l should be computed by the following equation:

$$R_{\text{in}}^l = \sum_{\mathbf{x}_i \in X_i} (s(\mathbf{x}_i, \mathbf{x}_{A(i)}) - s(\mathbf{x}_i, \mathbf{x}_l)). \quad (18)$$

Then, we take the greater value of 0 and R_{in}^l as the ultimate energy reduction for this case, as defined in the following equation:

$$R_{\text{in}} = \max\{0, R_{\text{in}}^l\}. \quad (19)$$

Namely, if R_{in}^l is the ultimate energy reduction, $\mathbf{x}_i \in X_i$ change their exemplars to $A(i)$. Otherwise, the current exemplar set is convincing and remains unchanged.

3.3.2. Case II. \mathbf{x}_l is not a current exemplar.

In this case, we define the current exemplar of \mathbf{x}_l as \mathbf{x}_e . When optimizing this situation, we firstly pretend to consider \mathbf{x}_l as a new alternative exemplar; namely, $E' = E, E'(\mathbf{x}_l) = l$. Then, similar to the analysis in case I, whether to accept \mathbf{x}_l as ultimate exemplar is decided by the reduction values of energy function. In detail, if \mathbf{x}_l is accepted as a feasible exemplar, some samples would change their exemplar from \mathbf{x}_e to \mathbf{x}_l . These samples are defined as $X_e^{l,e,l} = \{\mathbf{x}_i | s(\mathbf{x}_i, \mathbf{x}_l) > s(\mathbf{x}_i, \mathbf{x}_e)\}$. Thus, the corresponding energy reduction is defined as follows:

$$R_{\text{out}}^l = \sum_{\mathbf{x}_i \in X_e^{l,e,l}} (s(\mathbf{x}_i, \mathbf{x}_l) - s(\mathbf{x}_i, \mathbf{x}_e)). \quad (20)$$

On the other hand, may be current exemplar set E' is not convincing, so all samples would be certain to redetermine their exemplars including \mathbf{x}_l . As discussed before, the new exemplars for these samples are defined by equation (17), and the resulting energy reduction is listed as follows:

$$R_{\text{out}}^e = \sum_{\mathbf{x}_i \in X_e} (s(\mathbf{x}_i, \mathbf{x}_{A(i)}) - s(\mathbf{x}_i, \mathbf{x}_e)). \quad (21)$$

Remember that equations (20) and (21) are based on the assumption that $E' = E, E'(\mathbf{x}_l) = l$. Considering this, the energy reduction caused by \mathbf{x}_l which is not a current exemplar should be

$$R_{\text{out}} = s(\mathbf{x}_l, \mathbf{x}_l) - s(\mathbf{x}_l, \mathbf{x}_e) + \max_{l \in E} \{R_{\text{out}}^e, R_{\text{out}}^l\}. \quad (22)$$

To sum up, the optimization mechanism is shown below in detail.

3.4. Model Complexity. The similarity matrix is calculated according to the Euclidean distance; $s(\mathbf{x}_i, \mathbf{x}_j) = -\|\mathbf{x}_i - \mathbf{x}_j\|^2$. So, the scale of the similarity matrix is N^2 ; note that the amount of target data is not big. In the optimization process, we directly utilize the α -expansion move, which has $O(N^2)$ time complexity. For the proposed TrEEM, source-data-based exemplar set is actually one of the inputs and is out of the scope of the time complexity analysis of TrEEM here. Although we adopt EEM to obtain the source-data-based exemplar set $L(s)$, many other clustering models could be

helpful. TrEEM needs to select $L_s(i)$ from the source-data-based exemplar set in the first step, and this procedure has the time complexity of $O(N)$. In summary, the time complexity of TrEEM is $O(N^2)$ overall. Compared with other state-of-the-art transfer learning frameworks, it is very acceptable.

4. Experimental Results

To comprehensively evaluate the TrEEM model, we have conducted several experiments based on both synthetic and real-world datasets. For comparison, we also perform comparison with other different machine learning mechanisms, namely, EEM [15], multiclass SVM [23], TSK fuzzy system [24], and TSC [25] in the experiments. In this section, we will carefully analyze these experimental results.

4.1. Preparation. Before inputting the TrEEM model, we need to preprocess the original nonstationary EEG signals [1–3]. Usually, the features of EEG signals include time-domain features, frequency-domain features, and time-frequency features. In short, in time-domain analysis, statistics component features of the original EEG signals will be analyzed [26]. In frequency-domain analysis, power spectrum analysis and Short-Time Fourier Transforms (STFT) [27, 28] are commonly used. In time-frequency analysis, time domain and frequency domain are simultaneously extracted from high-dimensional and nonlinear EEG signals.

Various methods have been commonly used to extract EEG signals' features, including wavelet [29, 30], KPCA (Kernel Principal Component Analysis) [1, 2], and LDA (Linear Discriminant Analysis). In line with the experiments setting in [1–3], we use two feature extraction methods in this section, that is, KPCA and wavelet.

Besides, we use both synthetic and real-world datasets in this section. Firstly, we randomly generate 300 two-dimensional data points as 3 classes, shown in Figure 3. Then, we also choose Bonn EEG dataset [1, 2] as real-world data. The Bonn dataset is from the University of Bonn, Germany (<http://epileptologie-bonn.de/cms/upload/workgroup/lehnertz/eedata.html>), and has five classes. Each class (A to E) contains 100 signal channel EEG segments of 23.6 s duration. The sampling rate of all the datasets was 173.6 Hz. Each sample has 6 attributes. Table 3 lists a brief description of this dataset.

In addition, we examine the involved experimental results from two performance indices, namely, RandIndex (RI) [2, 31] and Purity. Assume that N is the total number of data points; we give the definitions of them below. That is, RI is shown in the following equation:

$$\text{RI} = \frac{f_{00} + f_{11}}{(N(N-1)/2)}, \quad (23)$$

where f_{00} is the amount of data whose cluster is in line with their class, while f_{11} is the amount of those data whose cluster is inconsistent with their class. Also Purity is defined in the following equation:

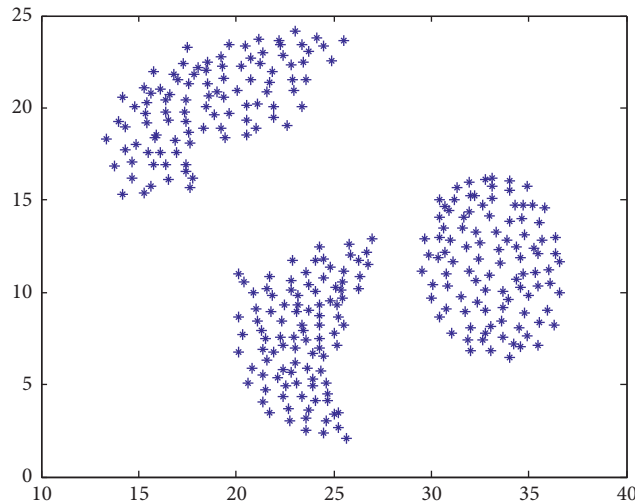


FIGURE 3: Synthetic dataset.

TABLE 3: Description of Bonn EEG dataset.

Subjects	Groups	Descriptions
Healthy	A	Signals captured from volunteers with eyes open
	B	Signals captured from volunteers with eyes closed
	C	Signals captured from volunteers during seizure silence intervals
Epileptic	D	Signals captured from volunteers during seizure silence intervals
	E	Signals captured from volunteers during seizure activity

$$\text{Purity}(E, C) = \frac{1}{N} \sum_k \max_j |e_k \cap c_j|, \quad (24)$$

where $E = \{e_1, \dots, e_N\}$ is the cluster result obtained by the learning model, while $C = \{c_1, \dots, c_N\}$ is the real data label set.

In all, the experiments are implemented in 2010a Matlab on a PC with 64-bit Microsoft Window 10, an Intel (R) Core (TM) i7-4712MQ, and 8 GB memory.

4.2. Results Analysis. As mentioned before, four machine learning methods are involved in this section, namely, EEM, multiclass SVM, TSK-FS, and the proposed TrEEM algorithm. There is no need to preset the cluster number in advance for EEM and TrEEM. In fact, it is a huge advantage for all exemplar-based clustering frameworks, whereas cluster number is an important parameter for TSK-FS. Multiclass SVM and TSC are two typical classification methods. Both EEM and TrEEM need parameter self-similarity $d(\mathbf{x}_i, \mathbf{x}_i)$. For multiclass SVM, in line with [23, 32], we choose Gaussian kernel function. In TSK-FS, usually the number of clusters is set to be equal to the number of fuzzy rules. Also, TSC need to preset the number of clusters. We follow the parameter setting strategy in relevant papers here. Besides, 5-fold cross validation is used to search the optimal parameters; and Table 4 lists brief introductions of these involved methods and the parameter searching range.

To construct the transfer learning scenario, for both synthetic and real-world EEG signal datasets, we randomly choose 80% data as source data and the remaining 20% as target data. For statistical analysis, in the experiment procedure, each algorithm is repeatedly executed 10 times; and we record the average performance and the corresponding standard deviation of RI and Purity. Furthermore, to deeply observe different extraction methods of EEG signals, we use both KPCA and wavelet here. The detailed comparison in terms of RI and Purity of the proposed TrEEM model and other benchmark approaches is shown in Table 5.

Observing Table 5, in this experimental setting, especially considering the fact that Bonn EEG signal dataset has 6 attributes and 5 classes, the performance of TrEEM model is very promising. TrEEM model is capable of recognizing useful information from both synthetic and real-world EEG signal datasets. Moreover, compared with other benchmark machine learning models, the proposed approach TrEEM performs better in terms of RI and Purity in this scenario.

In the experiment procedure, we also find that parameter self-similarity $d(\mathbf{x}_i, \mathbf{x}_i)$ has important influence on the experimental results, especially on the obtained number of clusters. The finding is identified with other exemplar-based clustering models [2, 12–15], and the parameter selection method is also in line with these models. See Table 4; we multiply $d(\mathbf{x}_i, \mathbf{x}_i)$ with η , and the value of η is decided from $\{0.01, 0.1, 1, 10, 50, 100\}$. In detail, with large value of η , the TrEEM model would obtain a smaller number of clusters,

TABLE 4: Parameters settings of involved algorithms.

Algorithms	Parameter setting
EEM: a typical exemplar-based clustering model	Self-similarity $d(\mathbf{x}_i, \mathbf{x}_j) = \text{median value of similarities} \times \eta \eta \in \{0.01, 0.1, 1, 10, 50, 100\}$.
Multiclass SVM: a typical classification learning model	Kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-(\mathbf{x}_i - \mathbf{x}_j ^2 / 2\sigma^2))$, where $\sigma \in \{2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5\}$. Penalty parameter $C \in \{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1, 10^2, 10^3\}$.
TSK-FS: a widely used fuzzy-rule-based learning model	FCM [1] involved, the cluster number equals the number of fuzzy rules. Number of fuzzy rules $K \in \{5, 10, 15, 20, 25, 30\}$. Gaussian membership function $\sigma \in \{2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5\}$.
TSC: transfer spectral clustering model	Preset the cluster number $n \in \{5, 10, 15, 20, 25, 30\}$; for other parameter setting strategies, see [25].
TrEEM: the proposed transfer exemplar-based learning model	Self-similarity $d(\mathbf{x}_i, \mathbf{x}_j) = \text{median value of similarities} \times \eta \eta \in \{0.01, 0.1, 1, 10, 50, 100\}$. Regularization factor $\lambda \in \{0.01, 0.05, 0.1, 0.5, 0, 5, 10, 15, 20, 25, 30\}$. $\sigma \in \{2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5\}$.
TSC: transfer spectral clustering model	Preset the cluster number $n \in \{5, 10, 15, 20, 25, 30\}$; for other parameter setting strategies, see [25].

TABLE 5: Comparison results of both synthetic and Bonn EEG datasets (the number in parentheses is the standard deviation).

Datasets (source data, target data, attributes, and classes)	Algorithms	Performance indices	
		RI	Purity
Synthetic dataset (240, 60, 2, 3)	EEM	0.8316 (0.0258)	0.7150 (0.1131)
	Multiclass SVM	0.8513 (0.0214)	0.8523 (0.0812)
	TSK-FS	0.8712 (0.0145)	0.8816 (0.0914)
	TSC	0.88230.0313	0.92360.0158
	TrEEM	0.8957 (0.0264)	0.9856 (0.0000)
Bonn EEG dataset (400, 100, 6, 5) (use KPCA to extract feature)	EEM	0.7754 (0.2146)	0.9800 (0.0000)
	Multiclass SVM	0.7827 (0.1834)	0.9643 (0.0023)
	TSK-FS	0.6819 (0.1579)	0.9623 (0.0000)
	TSC	0.72170.1241	0.9636 (0.0002)
	TrEEM	0.8323 (0.1652)	0.9600 (0.0012)
Bonn EEG dataset (400, 100, 6, 5) (use wavelet to extract feature)	EEM	0.7925 (0.0091)	0.7530 (0.0514)
	Multiclass SVM	0.7815 (0.0165)	0.9034 (0.0135)
	TSK-FS	0.7303 (0.0251)	0.9800 (0.0000)
	TSC	0.75010.1252	0.9600 (0.0021)
	TrEEM	0.8071 (0.0078)	0.9800 (0.0000)

while small η will bring in big cluster numbers. To fit with real data labels, in our experiments here, we set $\eta = 1$.

The regularization factor λ has a big effect as well. As analyzed before, λ determines how the source data affects the clustering result, and the value should not be too large. Obviously, if λ is too large, the clustering result based on the target data will be very close to that based on the source data, which is not what we want. When $\lambda = 0$, it means that TrEEM does not take the source data into account and TrEEM degrades to the typical EEM framework. In particular, Figures 4–6 show the dependence of model results on the value of λ . When $\lambda > 0$, in terms of RI and Purity, source data improves the performance of TrEEM. Index Purity is more sensitive to λ , while RI changes slowly.

Table 6 shows the average running time of 10 times for each approach. Yet the time consumption of the proposed TrEEM model is a bit more than those of EEM and TSK-FS; it is still in the same magnitude. Considering the improvements in RI and Purity, we think that the time complexity is acceptable. The results also fit the discussion in Section 3.4.

Therefore, from experimental results in Tables 5 and 6 and the above analysis, we can conclude the following:

- (1) For both synthetic and real-world EEG signal datasets, TrEEM performs great. Thus, we believe that TrEEM can effectively absorb knowledge from scant target data when similar source data exists.
- (2) For time consumption, TrEEM takes source data into account, which will inevitably increase the time complexity. Remember that the scale of target data will not be big, and the time consumption is very acceptable especially when combined with the performance in Table 5.
- (3) Although TrEEM requires the most parameters shown in Table 6, λ and η have big effects. Observing Figures 4–6, the performance of RI and Purity depends more on the value of λ . Note that we can narrow the optimization range according to the discussion in Section 3. Thus, we believe that parameter setting would be easy.

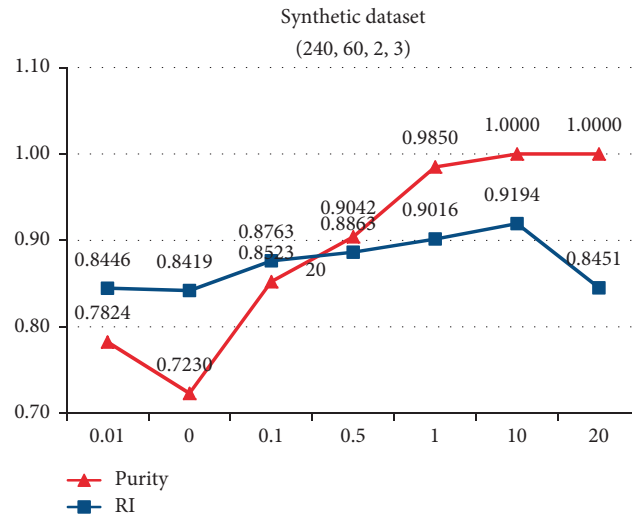


FIGURE 4: Effects of parameter λ on synthetic dataset.

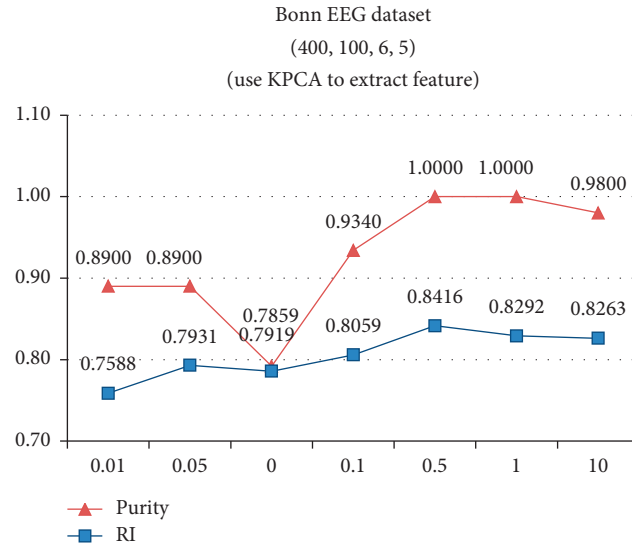


FIGURE 5: Effects of parameter λ on Bonn dataset with KPCA extract method.

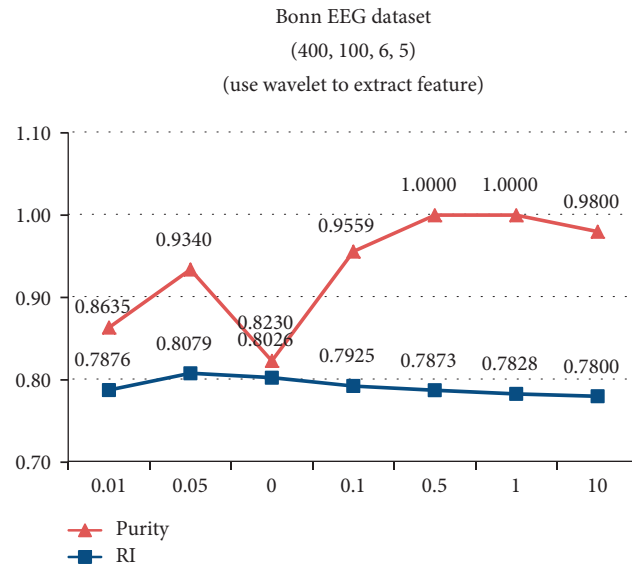


FIGURE 6: Effects of parameter λ on Bonn dataset with wavelet extract method.

TABLE 6: Average running time (Seconds) of the models on both synthetic and Bonn EEG datasets.

Datasets (source data, target data, attributes, and classes)	EEM	Multiclass SVM	TSK-FS	TSC	TrEEM
Synthetic dataset (240, 60, 2, 3)	0.1870	0.4562	0.0430	0.1520	0.2050
Bonn EEG dataset (400, 100, 6, 5) (use KPCA to extract feature)	0.6650	0.8934	0.3730	0.7923	0.6700
Bonn EEG dataset (400, 100, 6, 5) (use wavelet to extract feature)	0.4850	0.6327	0.1800	0.5331	0.5290

5. Conclusion

In conclusion, the contribution of this paper is providing a novel TrEEM framework to learn from few EEG signals when recognizing epileptic seizure. Starting from information theory, the proposed TrEEM method implants the similarity relationship between source and target data into the exemplar-based clustering model to improve the utilization rate of EEG signals, whereas this structure keeps all merits of the original optimization scheme. Therefore, without increasing the complexity of the model, TrEEM utilizes transfer learning method to learn from scant EEG signals. Yet our experimental results have shown promising performance of TrEEM, and several other perspectives should be considered as well. For instance, when each class contains unbalanced data, will this TrEEM method still work? And if we can provide multiple source data, what should we do to make them collaborate instead of bringing a negative effect? These are the problems that we should discuss in the future.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the 2018 Natural Science Foundation of Jiangsu Higher Education Institutions under Grant 18KJB5200001, by the Natural Science Foundation of Jiangsu Province under Grant no. BK20161268, and by the Humanities and Social Sciences Foundation of the Ministry of Education under Grant 18YJCZH229.

References

- [1] Y. Jiang, D. Wu, Z. Deng et al., "Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2270–2284, 2017.
- [2] Z. L. Bi Anqi and W. Ying, "Fast enhanced exemplar-based clustering for incomplete eeg signals," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 4147807, 11 pages, 2020.
- [3] Z. Jiang, F.-L. Chung, and S. Wang, "Recognition of multiclass epileptic eeg signals based on knowledge and label space inductive transfer," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 630–642, 2019.
- [4] Y. Jiang, Z. Deng, K.-S. Choi, F.-L. Chung, and S. Wang, "A novel multi-task TSK fuzzy classifier and its enhanced version for labeling-risk-aware multi-task classification," *Information Sciences*, vol. 357, pp. 39–60, 2016.
- [5] X. Zhu Jiaqi and L. Kang, "A novel double-index-constrained, multi-view, fuzzy-clustering algorithm and its application for detecting epilepsy electroencephalogram signals," *IEEE AC-CSS*, vol. 7, pp. 103 823–103 832, 2019.
- [6] D. Wen, P. Li, Y. Zhou et al., "Feature classification method of resting-state eeg signals from amnesic mild cognitive impairment with type 2 diabetes mellitus based on multi-view convolutional neural network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 8, pp. 1702–1709, 2020.
- [7] W. J. Xu Jiacan and H. Zheng, "Recognition of eeg signal motor imagery intention based on deep multi-view feature learning," *Sensors*, vol. 20, no. 12, 2020.
- [8] H. W. Liu ChienLiang and B. Xiao, "Epileptic seizure prediction with multi-view convolutional neural networks," *IEEE Access*, vol. 7, pp. 170 352–170 361, 2019.
- [9] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [10] L. L. Peng Yong, "Immune clonal algorithm based feature selection for epileptic eeg signal classification," in *Proceedings of the 2012 11th International Conference on Information Science, Signal Processing and Their Applications (ISSPA)*, Montreal, Canada, July 2012.
- [11] T. D. Z. Iscan and Z. Dokur, "Classification of electroencephalogram signals with combined time and frequency features," *Expert Systems With Applications*, vol. 38, no. 8, pp. 10499–10505, 2011.
- [12] S. W. Anqi Bi, "Incremental enhanced α -expansion move for large data: a probability regularization perspective," *International Journal of Machine Learning and Cybernetics*, vol. 8, no. 5, pp. 1–17, 2016.
- [13] K. Yizhang Jiang, "Anqi Bi, "Exemplar-based data stream clustering toward internet of things," *The Journal of Supercomputing*, vol. 76, no. 7, 2019.
- [14] A. Bi, F. Chung, S. Wang, Y. Jiang, and C. Huang, "Bayesian enhanced α -expansion move clustering with loose link constraints," *Neurocomputing*, vol. 194, pp. 288–300, 2016.
- [15] P. C. Y. Zheng, "Clustering based on enhanced α -expansion move," *IEEE Transaction on Knowledge and Data Engineering (TKDE)*, vol. 25, 2013.
- [16] W. T. F. M. F. Tappen, "Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters," in *Proceedings of the 9th IEEE International Conference on Computer Vision*, pp. 900–906, Nice, France, October 2003.
- [17] T. Kailath, "The divergence and bhattacharyya distance measures in signal selection," *IEEE Transactions on Communications*, vol. 15, no. 1, pp. 52–60, 1967.
- [18] T. M. Cover and T. Allen, *Elements of Information Theory*, Wiley Series in Telecommunications, Tsinghua University Press, Beijing, China, 1991.
- [19] Z. Liao, X. Lu, T. Yang, and H. Wang, "Missing data imputation: a fuzzy k-means clustering algorithm over sliding

- window,” in *Proceedings of the International Conference on Fuzzy Systems and Knowledge Discovery*, Tianjin, China, August 2009.
- [20] J. Zhang, Z. Yin, and R. Wang, “Pattern classification of instantaneous cognitive task-load through gmm clustering, laplacian eigenmap, and ensemble svms,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 14, no. 4, pp. 947–965, 2017.
- [21] Z. Yin and J. Zhang, “Identification of temporal variations in mental workload using locally-linear-embedding-based eeg feature reduction and support-vector-machine-based clustering and classification techniques,” *Computer Methods and Programs in Biomedicine*, vol. 115, no. 3, pp. 119–134, 2014.
- [22] A. P. Nguyen Chuong H. and K. Karavas George, “Inferring imagined speech using eeg signals: a new approach using riemannian manifold features,” *Journal of Neural Engineering*, vol. 15, no. 1, 2018.
- [23] V. N. Vapnik, “An overview of statistical learning theory,” *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [24] Z. Deng, Y. Jiang, K.-S. Choi, F.-L. Chung, and S. Wang, “Knowledge-leverage-based tsf fuzzy system modeling,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 8, pp. 1200–1212, 2013.
- [25] W. J. Jiang Yizhang and Z. Deng, “Transfer generalized fuzzy c-means clustering algorithm with improved fuzzy partitions by leveraging knowledge,” *Pattern Recognition and Artificial Intelligence*, vol. 26, no. 10, pp. 975–984, 2013.
- [26] B. R. Greene, S. Faul, W. Marnane, G. Lightbody, I. Korotchikova, and G. Boylan, “A comparison of quantitative eeg features for neonatal seizure detection,” *Clinical Neurophysiology*, vol. 119, no. 6, pp. 1248–1261.
- [27] W. Che Wan Fadzal and Mansor, “Short-time fourier transform analysis of eeg signal from writing,” in *Proceedings of the 2012 IEEE 8th International Colloquium on Signal Processing and its Applications*, Malacca, Malaysia, March 2012.
- [28] B. A. Vivaldi, “Frequency domain analysis of sleep eeg for visualization and automated state detection,” in *Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, NY, USA, September 2006.
- [29] L. Z. Q. Zhang and H. Kawabata, “Eeg Analysis Using Fast Wavelet Transform,” in *Proceedings of the 2000 IEEE International Conference on Systems, Man and Cybernetics. “Cybernetics Evolving to Systems, Humans, Organizations, and Their Complex Interactions*, pp. 2959–2964, Nashville, TN, USA, 2000.
- [30] X. C. D. Hu and W. Li, “Feature extraction of motor imagery eeg signals based on wavelet packet decomposition,” in *Proceedings of the 2011 IEEE/ICME International Conference on Complex Medical Engineering*, pp. 694–697, Harbin, China, May 2011.
- [31] Q. W. Cao F, “Density-based clustering over an evolving data stream with noise,” in *Proceedings of the 2006 SIAM International Conference on Data Mining*, pp. 328–339, SIAM, Bethesda, MD, USA, April 2006.
- [32] Z. J. Zhang Yuanpeng and X. Li, “Epileptic eeg signals recognition using a deep view-reduction tsf fuzzy system with high interpretability,” *IEEE Access*, vol. 7, no. 99, pp. 137 344–137 354, 2019.

Research Article

Automatic Retinal Vessel Segmentation Based on an Improved U-Net Approach

Zihe Huang ¹, Ying Fang ¹, He Huang ¹, Xiaomei Xu ¹, Jiwei Wang ²,
and Xiaobo Lai ¹

¹School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou, China

²Department of Information, The 73th Group Army Hospital of P. L. A, Xiamen, China

Correspondence should be addressed to Jiwei Wang; 772073190@qq.com and Xiaobo Lai; dmia_lab@zcmu.edu.cn

Received 13 February 2021; Revised 30 March 2021; Accepted 8 April 2021; Published 24 April 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Zih Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Retinal blood vessels are the only deep microvessels in the blood circulation system that can be observed directly and noninvasively, providing us with a means of observing vascular pathologies. Cardiovascular and cerebrovascular diseases, such as glaucoma and diabetes, can cause structural changes in the retinal microvascular network. Therefore, the study of effective retinal vessel segmentation methods is of great significance for the early diagnosis of cardiovascular diseases and the vascular network's quantitative results. This paper proposes an automatic retinal vessel segmentation method based on an improved U-Net network. Firstly, the image patches are rotated to amplify the image data, and then, the RGB fundus image is preprocessed by normalization. Secondly, after the improved U-Net model is constructed with 23 convolutional layers, 4 pooling layers, 4 upsampling layers, 2 dropout layers, and Squeeze and Excitation (SE) block, the extracted image patches are utilized for training the model. Finally, the fundus images are segmented through the trained model to achieve precise extraction of retinal blood vessels. According to experimental results, the accuracy of 0.9701, 0.9683, and 0.9698, sensitivity of 0.8011, 0.6329, and 0.7478, specificity of 0.9849, 0.9967, and 0.9895, F1-Score of 0.8099, 0.8049, and 0.8013, and area under the curve (AUC) of 0.8895, 0.8845, and 0.8686 were achieved on DRIVE, STARE, and HRF databases, respectively, which is better than most classical algorithms.

1. Introduction

The retinal blood vessels are essential parts of the whole microcirculation system with extremely wealthy blood vessel information. They are the only deep microvessels in the blood circulation system that can be observed directly and non-invasively. Most ophthalmic diseases are attributed to fundus retinopathy [1], glaucoma, diabetic retinopathy, and age-related macular degeneration, the four primary reasons. Therefore, early detection, diagnosis, and treatment can effectively reduce the degree of patients affected by the disease or slow down its progress. The retinal blood vessels are also the main anatomical structures visible from color fundus images [2]. Color fundus images can directly detect retinal vascular microaneurysms, fundus hemorrhage, exudate, and other diseases. They are essential bases for ophthalmologists to diagnose and treat diseases. Accurate segmented images play an essential role in the preliminary screening, subsequent

diagnosis, and treatment of fundus diseases. The manual segmentation is tedious, time consuming, and requires skilled techniques. When different ophthalmologists segment the same blood vessel image in clinical practice, the results obtained are different, which will cause errors in the final segmentation results [3]. The automatic segmentation technology of retinal vessels can relieve ophthalmologists' diagnostic pressure and effectively solve inexperienced ophthalmologists' problems. Therefore, the automatic segmentation of retinal vessels is important for the clinical diagnosis and treatment of ophthalmic diseases.

The highly complex structure of retinal vessels, low contrast between the retinal target vessel and the background image, and noise interference when acquiring retinal images can significantly increase the difficulty of fundus retinal blood vessel segmentation. Although there are many barriers to realizing the accurate segmentation of retinal vessels, it is of considerable significance to auxiliary clinical

diagnosis and treatment. In recent years, it has attracted many scientific experts to explore the field and achieved fruitful results [4]. According to the learning mode, the existing retinal vascular segmentation methods can be roughly classified as unsupervised methods [5–16] and supervised [17–19].

The unsupervised method can be subdivided into different methods, such as matched filtering, vascular tracking, morphological processing, and model-based approaches. Fraz et al. [5] combined vascular centerline detection with morphological bit plane section technology to extract blood vessels from retinal images. To detect the blood-vessel-like pattern in a noisy environment, Zana and Klein [6] implemented an algorithm about mathematical morphology and curvature evaluation. Niemeijer et al. compared many vascular segmentation algorithms [7]. According to the research, the highest accuracy of these algorithms is 0.9416. Mendonça and Campilho [8] combined the strength and partial morphological features of the vascular structure with the different filters for centerline extraction and the morphological operator for filling vascular segments to gain the final segmentation results of fundus retinal vessels. Garg et al. [9] proposed a curvature-based method combined with an improved region growth method to extract complete vascular structures from color retinal images. Chutatape et al. [10] used Gaussian filtering and Kalman filtering to make the best linear estimation of subsequent blood vessels' position on the known centerline of blood vessels. Gang et al. [11] obtained a two-dimensional Gaussian filtering method through mathematical analysis and simulation experiment process, which significantly affects amplitude correction and improves the detection success rate. Vlachos and Dermatas [12] presented a retinal vessel image extraction algorithm whose basic principle is iterative line tracking and morphological postprocessing, with an average accuracy of 0.929. Hoover et al. put forward a threshold detection algorithm based on matched filter [13]. To eliminate the optic disc and background noise in the fundus image, Zardadi et al. [14] enhanced the blood vessels in all directions and used the adaptive threshold as an unsupervised classifier for pixel classification and perform morphological postprocessing at last. Srinidhi et al. [15] used local and global information to decompose the vascular tree into multiple subtrees and divide the vascular subtrees according to arteries and veins, respectively. Upadhyay et al. [16] implemented local directional-wavelet transform and global curvelet transforms to enhance vessel and segmentation. The supervised method uses the truth map as the training sample to train the model. Aslani and Sarnel come up with mixed feature vectors from different algorithms to represent pixels [17] and mixed feature vectors to train the random forest (RF) classifier to classify vascular/nonvascular pixels. Marin et al. employed a neural network scheme to classify pixels and calculate 7-dimensional vectors composed of grayscale and moment-invariant features to represent pixels [18]. Jiang et al. [19] proposed a D-Net model for retinal blood vessels, which can obtain denser feature information and reduce the excessive loss of tiny blood vessels' feature information.

Traditional supervision methods are usually defined by experience and require manual intervention, which may result in deviation. Therefore, an automatic and effective feature extractor deep learning is necessary to achieve higher efficiency. Wang et al. combined two advanced classifiers named the convolution neural network (CNN) and random forest (RF) for segmentation [20]. Maji et al. processed color fundus images and detected blood vessels on the foundation of the ConvNet integrated algorithm framework [21]. Che et al. utilized supervised blood vessel segmentation based on an artificial neural network to estimate the effect of aging on the performance [22]. The results illustrate that people of different ages have an impact on various aspects of the subdivision results. Liskowski and Krawiec presented deep neural network segmentation architecture with an extensive training dataset strengthened by global contrast normalization, zero-phase preprocessing, geometric transformation, and gamma correction [23]. Tan et al. took advantage of a 10-layer CNN segment and distinguished exudates, bleeding, and microaneurysms simultaneously [24]. The network structure consists of two stages to improve the performance, segment the fundus image's pathological features with superior accuracy, and demonstrate the strong feature extraction ability of deep learning. Wu et al. [25] integrated residual learning with DenseNet, which uses each layer of vascular feature map information effectively so that the model can obtain more robust morphological structure information according to the gold standard. However, due to the excessive use of dense connection modules, the amount of memory will be expanded, and the computational complexity will be added. Fu et al. [26] used a deep learning framework based on fully convolutional networks (FCN) to figure out the problem of vascular edge detection. Lin et al. [27] proposed an automatic retinal blood vessel segmentation network that combines globally nested edge detectors with global smoothness regularization from conditional random fields. Khan et al. [28] extended the FCN variant, which significantly reduced the number of adjustable hyperparameters and decreased the computational overhead of the training and testing phases.

Although the automatic segmentation algorithm of retinal vessels can effectively segment the retinal vascular network, the automatic segmentation method's performance still needs to be improved. There are some limitations, such as insufficient segmentation degree and poor continuity of microretinal vessels, which cannot meet clinical diagnosis needs. The U-Net model has been applied to the segmentation of medical images for years, and many improved structures based on the U-Net model [29–36] have achieved good segmentation results. In this paper, we explore an automatic segmentation algorithm of retinal vessels based on improved U-Net. The proposed scheme of the work is shown in Figure 1. It contains training and testing stages. In the training stage, the retinal vessel training set images and the gold standard images are preprocessed, and the image blocks are extracted to train the improved U-Net model. The parameters are learned by training data. In the test stage, the same preprocessing method is used for the test images, and the test set image blocks are tested and extracted through the training model.

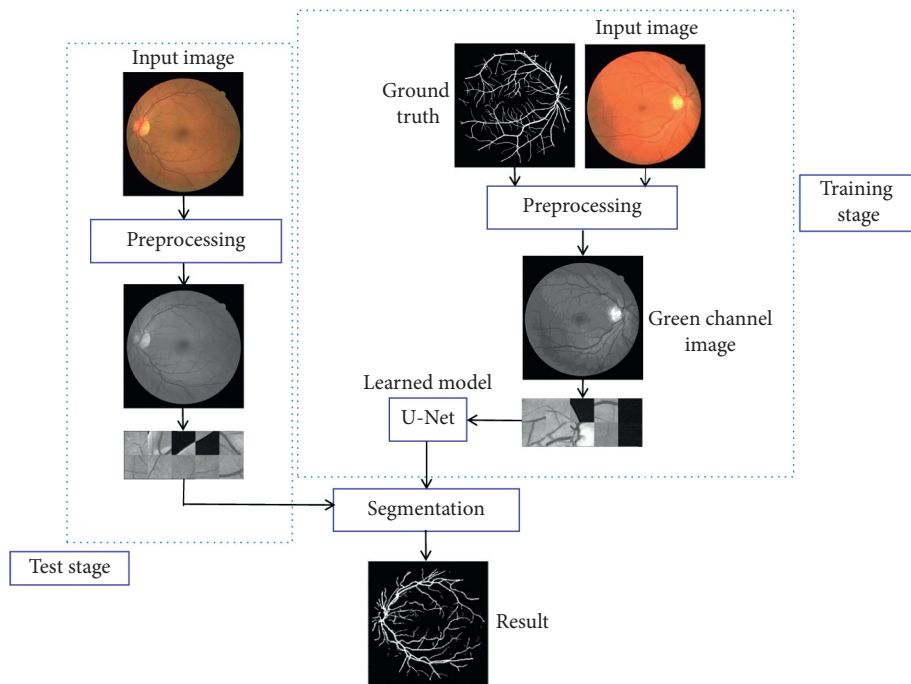


FIGURE 1: Overview of the proposed framework.

This article's remainder structure is arranged as follows: Section 2 introduces the improved automatic segmentation method of retinal vessels in detail. Section 3 will compare and analyze our experimental results. Finally, Section 4 draws conclusions about this work.

2. Methods

In order to improve algorithm performance, we combine partial advantages of the unsupervised and supervised methods. Firstly, the RGB fundus image is preprocessed. Secondly, the U-Net model is constructed with a SE block, and the extracted patches are applied to train the improved U-Net model. Finally, the trained model is tested to obtain the eventual results. The overview of the segmentation framework is shown in Figure 1, which consists of three steps:

Step 1: fundus image preprocessing: we extract the training set's patches and the gold standard of fundus image, and the patches are rotated to amplify the data and normalized subsequently

Step 2: building and training the model: we construct the U-Net model with an SE block and set up the parameters, and the extracted patches are applied to train the model

Step 3: testing and segmenting retinal vessels. Firstly, the test set's fundus patches are extracted and normalized; then, we test and extract retinal vessels automatically through the trained model

2.1. Materials. For the sake of evaluating the effectiveness of the algorithm, the fundus image datasets DRIVE (<http://www.isi.uu.nl/Research/Databases/DRIVE/>), STARE ([\[cecas.clemson.edu/~ahoover/stare/\]\(http://cecas.clemson.edu/~ahoover/stare/\)\), and HRF \(<https://www5.cs.fau.de/research/data/fundus-images/>\) are used to segment the retinal vessels. DRIVE is a fundus image dataset usually referred to for evaluating the performance of retinal vascular segmentation methods in other experiments. DRIVE is concluded from the Netherlands DR screening Project. It contains 40 color fundus images with a resolution of \$768 \times 584\$ pixels. The STARE dataset consists of 20 normal and diseased fundus images with \$605 \times 700\$ pixels. The full name of HRF is High-Resolution Fundus Image Database. HRF consists of 15 images of healthy patients, 15 images of patients with diabetic retinopathy and 15 images of patients with glaucoma, all collected by using clinical fundus cameras with a resolution of 3504 pixels.](http://</p>
</div>
<div data-bbox=)

2.2. Preprocessing. Our model is prone to overfitting during the training process due to the relatively small number of training samples in our DRIVE, STARE, and HRF databases. Simultaneously, the collected retinal images are susceptible to mutual interference. Preprocessing that contains the amplification and normalization of the fundus image data is necessary to deal with these problems. Figure 2 shows the results of the image normalization, where (a) is the original image and (b) is the normalized image. The process is briefly described as follows:

- (1) Image data amplification: patch extraction and geometric transformation are used to amplify the image data.
 - ① Taking each voxel as the center, the two-dimensional region with the size of $n \times n$ pixels is extracted as the corresponding patches of the voxel. The window size of the model input is 48×48 .
 - ② The images of the training set and the patches extracted by the gold standard are rotated. Figure 2,

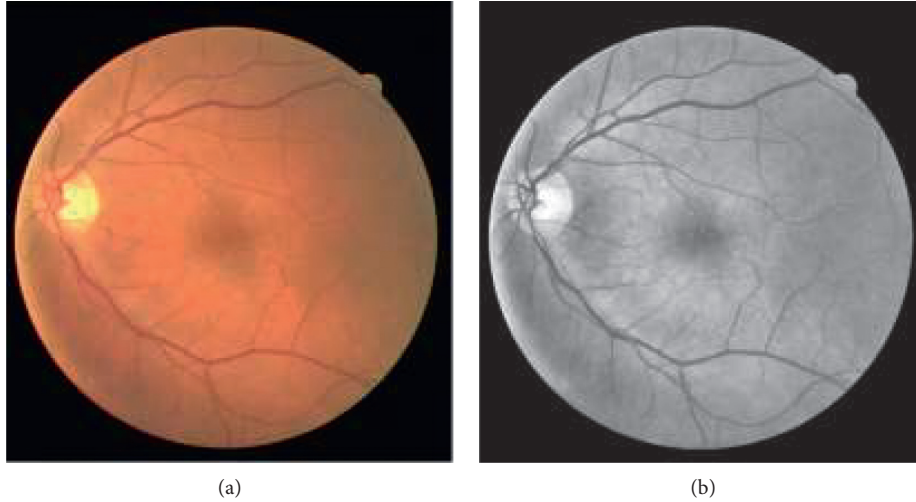


FIGURE 2: Image normalization.

(a) is the extracted image patches, and Figures 3(b)~3(d) show visualizations of the image patches rotated by 90° , 180° , and 270° , respectively.

- (2) Image normalization: there is inevitable noise interference in the fundus image. The features are normalized so that they have a mean of 0 and a variance of 1. The mean and variance are standardized as follows:

$$x_k^* = \frac{x_k - \mu}{\sigma}, \quad (1)$$

where x_k denotes the value of the k_{th} pixel before standardization, x_k^* represents the intensity of the k_{th} voxel after standardization, μ is the average intensity of all pixels, and σ is the standard deviation of the intensity of all pixels.

2.3. Improved U-Net Architecture Model. The U-Net model was first put forward by Ronneberger et al. [37] in 2015 and applied to the cell segmentation of medical images. In recent years, Alom et al. [38] applied U-Net to retinal vascular segmentation, combined with a cyclic neural network, to further improve retinal vascular segmentation accuracy. The U-Net neural network is an end-to-end training supervised image segmentation network based on FCN. It has excellent performance under fewer datasets and can be widely used in medical image segmentation research. This paper optimizes the basic structure of U-Net.

2.3.1. SE Block. SE block is a network substructure in SENet. It can be easily integrated into almost any network architecture. As for the traditional convolution process, its channel dependence is not fully utilized. Moreover, the filters learned by the channel are all performed in the local receptive field, resulting in each feature map being unable to use the context information outside the region. The channel dimension of the SE block in the convolutional layer adds

the attention mechanism. It aims to adopt a feature recalibration strategy to learn the importance of each feature from the global information through the network loss function. Then, the valuable features are promoted according to their importance, and the less valuable features are suppressed. This training method improves the performance of the network. The addition of the SE block will only add a small amount of calculation and parameters to the original network, which can be wholly accepted compared with the improvement of its effect.

SE block is roughly divided into 3 parts. For an input x , supposing the feature channels value is $c1$, it would get a particular feature after some transformations such as convolution. Moreover, this particular feature has the value of feature channels $c2$. Unlike the traditional CNN, the features previously obtained need to be recalibrated next. Firstly, in the squeeze stage, the feature map should go through a process which is named global average pooling. The purpose of this process is to convert two-dimensional feature channels into an actual number. In a sense, there is a global receptive field for each actual number so that the low convolutional layer could also use information obtained by the global receptive field. Secondly, the next stage is the excitation operation. It is analogous to the gate mechanism in the recurrent neural network. Each feature channel should get a weight identified by the parameter W , and the parameter W is regarded as the correlation between feature channels. Finally, the reweight operation always recounts a weight as the importance of each feature channel. Usually, it is the weight of the output of the excitation and then weights the previous feature channel by channel through multiplication to achieve the original feature in the channel dimension recalibration operation.

2.3.2. Improved U-Net Network Architecture. The U-Net network model has an encoder and a decoder, and it includes two parts: the contraction path on the left and the expansion path over there. As Figure 4 shows, the network

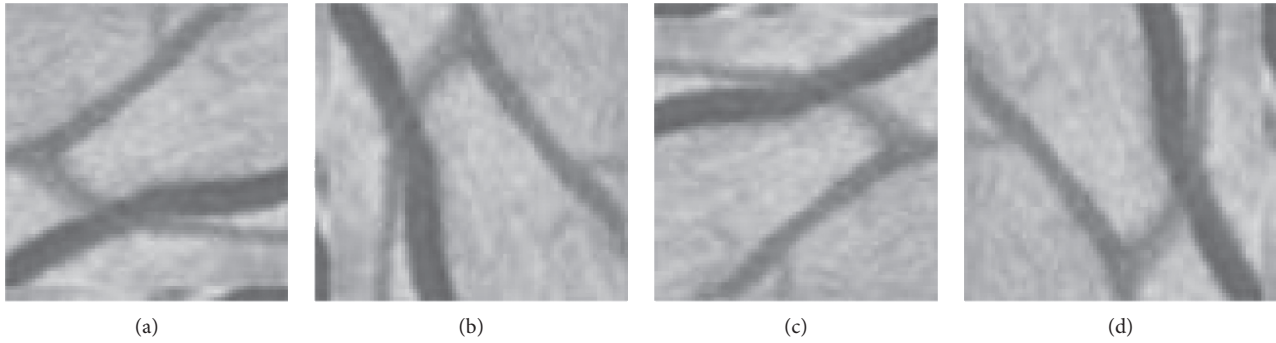


FIGURE 3: Image data augmentation. (a) Original image patch, (b) rotate 90°, (c) rotate 180°, and (d) rotate 270°.

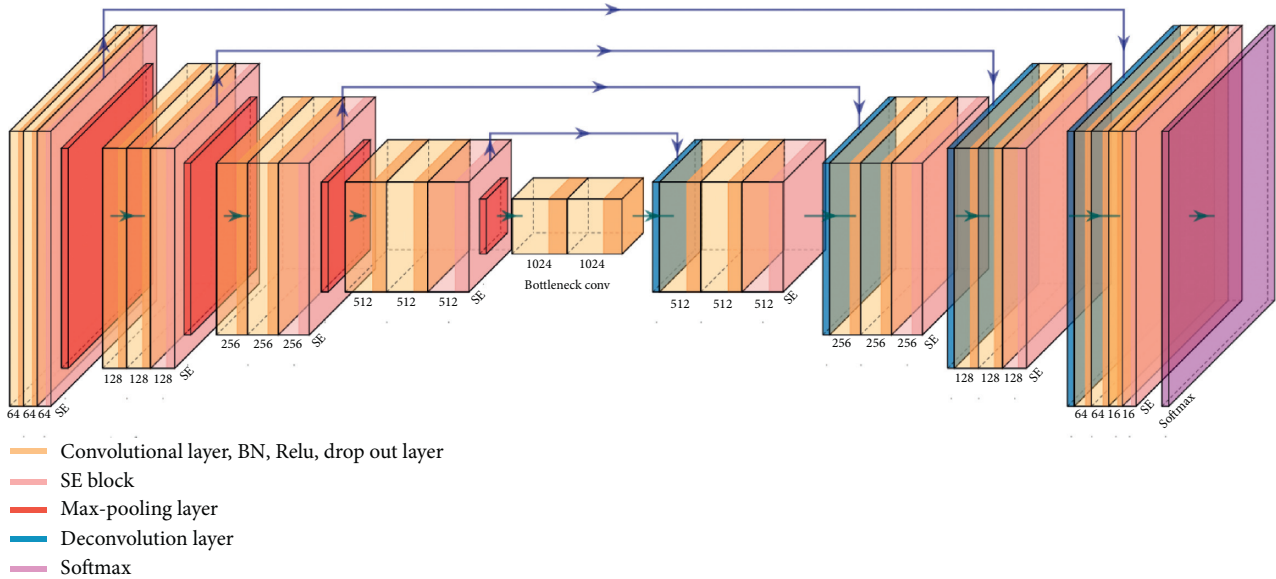


FIGURE 4: Improved U-net network architecture.

contains a 3×3 convolutional layer, the copy and cropping, and the max-pooling layer. The deconvolution layer has a step size of two. Rectified linear unit (Relu) is a nonlinear activation function used for the entire network structure. According to the feature recalibration, the network could embed the SE block. Then, the interdependence between feature channels is modelled explicitly. It is worth noting that each feature channel would be obtained automatically according to the learning stage. Then, the block promotes useful features and suppresses the less valuable features for the current task.

The contraction path of the network is an encoder that adopts the structure of a convolutional neural network. Two 3×3 convolution layers are used for feature extraction, and the activation function of Relu is applied to improve the U-net network's expression ability. Next is the channel attention mechanism used by the 1-layer SE module to add feature maps. Then, a 2×2 max-pooling layer is used to reduce the resolution of the input signal. The extension path is the decoder, which restores the image size through upsampling. It improves the subdivision capability and contour prediction level of the model through the data

features extracted by the joint downsampling process (the visualization result of downsampling and upsampling is exhibited in Figure 5). Finally, a 1×1 convolutional layer is applied to convert the final output channel number into the number of divided categories.

Our model mainly sets up 23 convolution layers, 4 pooling layers, 8 SE layers, 4 upsampling layers, and 2 dropout layers. Each layer's structure consists of an input layer, 4 contraction paths, a bottom layer, 4 extension paths, and an output layer. The activation layer is the Relu activation function.

2.3.3. Training Improved the U-Net Model. To make the experiment more rigorous, the segmentation results of retinal vessels should be kept accurate and stable. Compared with the traditional U-Net, batch standardization (BN) is applied to improve the training stage's stability and reduce the possibility of gradient disappearance. The segmentation performance is optimized, and the effect has remarkable universality. Besides, BN helps the model converge. The rational formula is as follows:

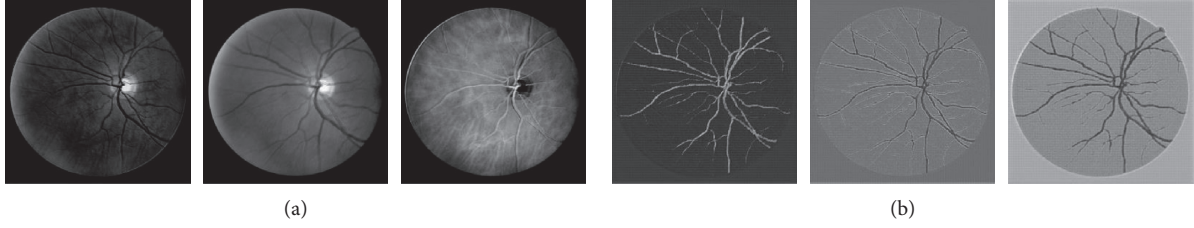


FIGURE 5: Visualization result of downsampling and upsampling. (a) Downsampling. (b) Upsampling.

$$y = \frac{\gamma}{\sqrt{\text{Var}[x] + \varepsilon}} \cdot x + \left(\beta - \frac{\gamma \cdot E[x]}{\sqrt{\text{Var}[x] + \varepsilon}} \right), \quad (2)$$

where x represents the input feature, y indicates the standardized feature, ε is close to zero, and γ and β denote the training parameters, which can be iteratively updated during the training process.

The cross-entropy loss function is applied in the process of model training. The optimization of the parameters of the whole network would be achieved owing to the Adam algorithm. The parameter update process of the algorithm is as follows:

$$\begin{aligned} x_n &= x_0 - \frac{\eta m}{\sqrt{v_n}}, \\ m_n &= a_1 \times m_0 + (1 - a_1) f'(x_0), \\ v_n &= a_2 \times v_0 + (1 - a_2) [f'(x_0)]^2, \\ v_n &= a_2 \times v_0 + (1 - a_2) [f'(x_0)]^2, \end{aligned} \quad (3)$$

where a_1, a_2 represent the loss rate, η is the learning rate, v_n, v_0 denote the new and old parameters, and m represents the momentum. The algorithm can determine the range of learning rates in an iteration to ensure the stability of parameters and high computational efficiency. In the model, the epoch is set to 10, and an iteration (epoch) is a complete transmission of all training samples. The algorithm learning rate is adjusted to every 5 epochs, and the learning rate is reduced to 0.96. At the same time, the dropout is set to 0.5. Half of the training data are randomly discarded to prevent overfitting.

3. Results and Discussion

3.1. Parameter Configuration. 10 iterations can achieve the ideal training accuracy in this paper. By optimizing the cross-entropy loss function, the pixel segmentation error rate is minimized. This paper uses the Adam algorithm to optimize the loss function, and the learning rate is set to 0.001. The cross-entropy function is defined as follows:

$$H(p, q) = - \sum_x p(x) \log q(x). \quad (4)$$

The cross entropy represents the distance between the two probability distributions. The difficulty of the probability distribution $p(x)$ is represented by the probability

distribution $q(x)$, in which p is the correct answer and q is the predicted value. The smaller the cross entropy is, the closer the distribution of the two probabilities. The software applied in this experiment are Pycharm (python3.6), Keras, and its TensorFlow port. The running processor is Intel (R) Core (TM) i7-7700HQ CPU @ 2.81 GHz, with 16 GB memory. The GPU graphics card is NVIDIA GeForce GTX 1050, and the operating system is 64 bit Windows 10.

3.2. Evaluation Metrics. We use three commonly used metrics to evaluate the performance objectively, including accuracy (ACC), sensitivity (TPR), specificity (TNR), F1-score, and area under the curve (AUC). Different metrics are calculated as follows:

$$\begin{aligned} \text{ACC} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}, \\ \text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{TNR} &= \frac{\text{TN}}{\text{FP} + \text{TN}}, \\ \text{F1-score} &= \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}}, \end{aligned} \quad (5)$$

where ACC represents the proportion of the sum of the correctly segmented pixels to the sum of all pixels, TPR is the ratio of the correctly segmented retinal vessel pixels to the pixels in the gold standard image, and TNR represents the proportion of the correctly segmented background pixels to the real background. TP in the formula is true positive, indicating the number of retinal vessels correctly segmented. TN, which represents the number of background pixels correctly segmented, is a true negative value. FP means false positive, suggesting that the number of pixels segmented into blood vessels by mistake. FN is a false negative, showing the number of pixels erroneously segmented into the background. AUC represents the area under the ROC curve and is also adopted as an essential evaluation metric for vessel segmentation. The values of AUC and F1-score are 1 for a perfect classifier. Figure 6 shows ROC curves of the proposed model training and testing on DRIVE, STARE, and HRF.

3.3. Results. The improved U-Net model is used to extract retinal vessels from the databases of DRIVE, STARE, and HRF color fundus images. The test results are shown in

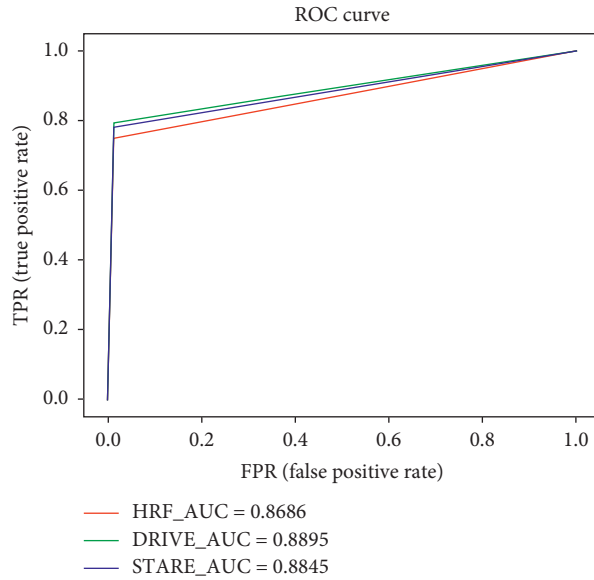


FIGURE 6: ROC curves of the proposed model training and testing on DRIVE, STARE, and HRF.

Figures 7–9, representing the segmentation results of retinal vessels from DRIVE, STARE, and HRF databases. The first column is the original image, and the second column indicates the improved U-Net segmentation result; then, the third column represents the result of expert manual segmentation.

Table 1 shows the quantitative results of ACC, TPR, and TNR obtained by automatic segmentation of retinal vessels from the DRIVE fundus image database. Table 2 exhibits the segmentation of 20 images from the STARE database. Table 3 shows the average values (AVG) and standard deviation (STD) of ACC, TPR, and TNR from the DRIVE, STARE, and HRF color fundus image databases.

3.4. Discussion. The segmentation results using the improved U-Net model are visually compared with the segmentation results of Zhang, Jiang, and Zana's three algorithms. It can be seen from the segmentation diagram in Figure 10 that the proposed algorithm has better results than the other three algorithms. Compared with Jiang and Zhang's algorithm, our algorithm has higher integrity and a more significant advantage in vascular continuity. In the segmentation background, our segmentation effect is closer to the manual segmentation result than Zhang and Zana's algorithms. The vascular trunk and ending connectivity is excellent, but most of the microretinal vessels are better segmented. Therefore, we can provide a valuable reference for clinical diagnosis.

We compared our improved U-Net with several state-of-the-art methods on the HRF database, and the segmentation results are shown in Figure 11. It can be seen subjectively and objectively that AG-UNet [40] and M-GAN [41] have limitations in capturing those thin blood vessels. IterNet [42] seems to miss the vessels around the optic discs on images. NMF+3D U-Net [43] and our method have shown better performance, but our method is more sensitive to small

blood vessels and produces more distinct segmentation results. The proposed method is able to detect weak vessels that may be lost in AG-UNet and M-GAN. Thus, it is more powerful to preserve considerable details.

Tables 4 and 5 show the quantitative comparison of the ACC, TPR, and TNR results obtained by the proposed algorithm and partial unsupervised and supervised method for retinal vascular segmentation of DRIVE and STARE fundus image database images, respectively. In the experiment, the result of manual segmentation by the first expert in the DRIVE and STARE fundus image database is used as the ground truth. As Table 4 shows, this algorithm has achieved a higher ACC than most vascular segmentations, except for that of Wang et al. [20] and Khowaja et al. [52], whose methods are based on hierarchical classification. Wang introduced CNN and random forest into hierarchical classification. Thus, the model has strong generalization capability. Khowaja adopts the hierarchical feature extractor instead of the hybrid feature set with subspace learning methods to improve the segmentation performance. It is noteworthy that the TNR performance of this algorithm is excellent, and its performance is better than that of other algorithms. As for TPR, our algorithm is better than that of Zana and Klein [6], Lin et al. [27], and other unsupervised methods. The TPR of the generative adversarial network (GAN) is based on conditional patches as high as 0.7746, second only to the algorithm we proposed. GAN proposed by Khan et al. [28] utilizes a generator network and a patch-based discriminator network conditioned on the sample data with an additional loss function to learn both thin and thick vessels. The model is possible to probe the different patch sizes so that GAN has a certain level of competitiveness with current advanced techniques. Table 5 shows that although the number of vascular images in the training set is not large, our algorithm is still effective in the test set and can obtain ideal segmentation results with excellent robustness and generalization ability. It could be observed from Table 6 that

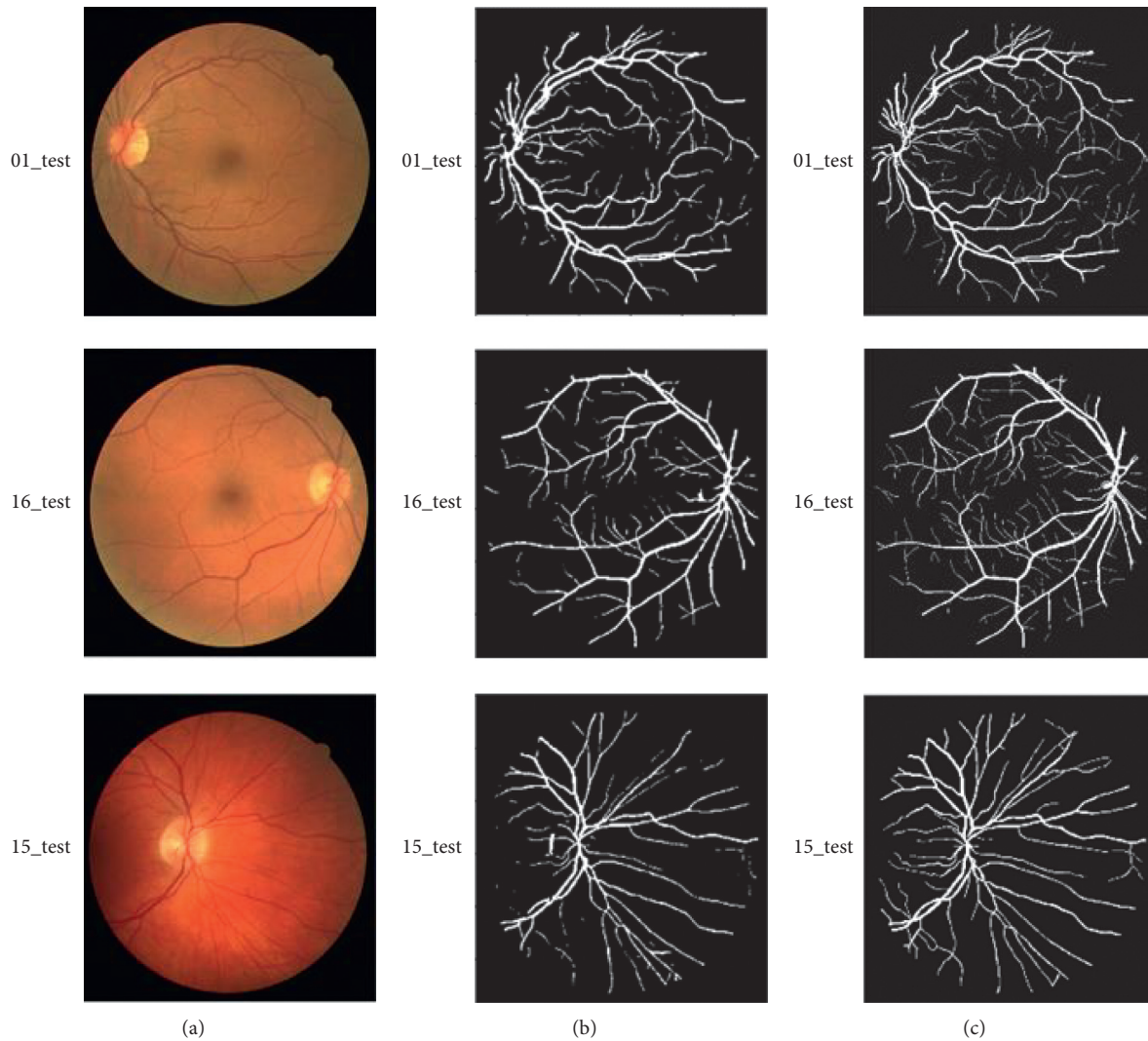


FIGURE 7: Visualization of the random samples results for the DRIVE image dataset. (a) Original image, (b) U-Net, and (c) manual.

all methods achieve a similar level of specificity where M-GAN is 0.0036 and NMF+3D U-Net is 0.0008, slightly higher than ours. Comparing to the listed methods, AG-UNet is slightly more sensitive under the HRF database while IterNet is second only to AG-UNet. Our approach and M-GAN both reach preferable performance where the accuracy of M-GAN is 0.0002 higher than ours.

In recent years, many improved structures based on the U-Net model have achieved good segmentation results. However, our improved model is based on the new attention mechanism module proposed in 2020cvpr. There is no article applying the Squeeze-Attention module to the field of medical image segmentation, so the improved U-Net model

we proposed still has a certain degree of innovation and novelty.

To prove the effectiveness of the proposed model, we compared the results of the algorithm with other state-of-the-art modified U-Net as follows. It can be noticed from Table 7 that our algorithm has an advantage over other modified U-Net models. However, the sensitivity of our method on the STARE database is the lowest. One possible reason is that the number of images from the STARE database is too small to achieve the original effect. Therefore, we can learn from other excellent segmentation methods to improve preprocessing and finally achieve the best results.

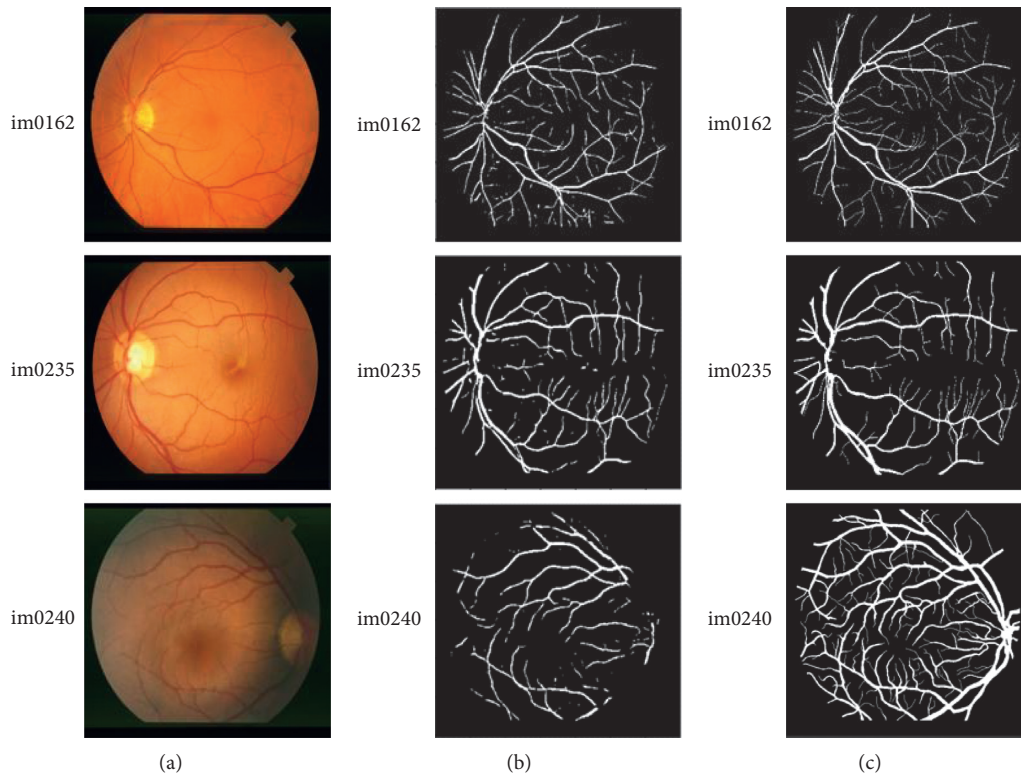


FIGURE 8: Visualization of the random samples results for the STARE image dataset. (a) Original image, (b) U-Net, and (c) manual.

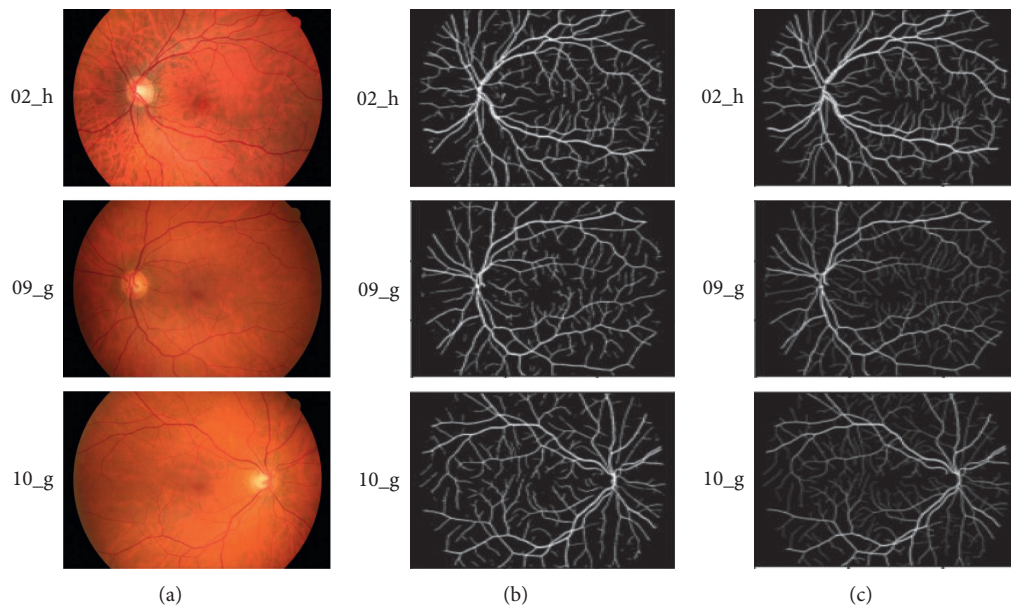


FIGURE 9: Visualization of the random samples results for the HRF image dataset. (a) Original image, (b) U-Net, and (c) manual.

TABLE 1: Quantitative results of the proposed method for the DRIVE fundus database.

Image	ACC	TPR	TNR
01	0.9646	0.8386	0.9779
02	0.9676	0.8001	0.9879
03	0.9577	0.6764	0.9900
04	0.9677	0.7663	0.9891

TABLE 1: Continued.

Image	ACC	TPR	TNR
05	0.9621	0.6464	0.9947
06	0.9601	0.6831	0.9911
07	0.9630	0.6973	0.9938
08	0.9578	0.6093	0.9943
09	0.9654	0.6630	0.9938
10	0.9680	0.7683	0.9877
11	0.9650	0.7594	0.9882
12	0.9676	0.7592	0.9901
13	0.9619	0.7703	0.9838
14	0.9690	0.7661	0.9886
15	0.9715	0.7547	0.9897
16	0.9668	0.7636	0.9880
17	0.9633	0.6837	0.9909
18	0.9673	0.8114	0.9816
19	0.9729	0.8923	0.9811
20	0.9692	0.8035	0.9831

TABLE 2: Quantitative results of the proposed method for the STARE fundus database.

Image	ACC	TPR	TNR
im0162	0.9691	0.8231	0.9811
im0163	0.9727	0.8659	0.9817
im0235	0.9691	0.8073	0.9849
im0236	0.9683	0.7868	0.9864
im0239	0.9515	0.5499	0.9914
im0240	0.9423	0.4981	0.9940
im0255	0.9648	0.6624	0.9955
im0291	0.9709	0.4812	0.9975
im0319	0.9748	0.4856	0.9973
im0324	0.9523	0.3371	0.9970

TABLE 3: AVG and STD of the evaluation index of the DRIVE, STARE, and HRF fundus database.

Database	ACC		TPR		TNR	
	AVG	STD	AVG	ATD	AVG	STD
DRIVE	0.9701	0.0287	0.8011	0.3817	0.9849	0.0194
STARE	0.9683	0.0315	0.6329	0.3987	0.9967	0.0208
HRF	0.9698	0.0231	0.7478	0.3029	0.9895	0.0115

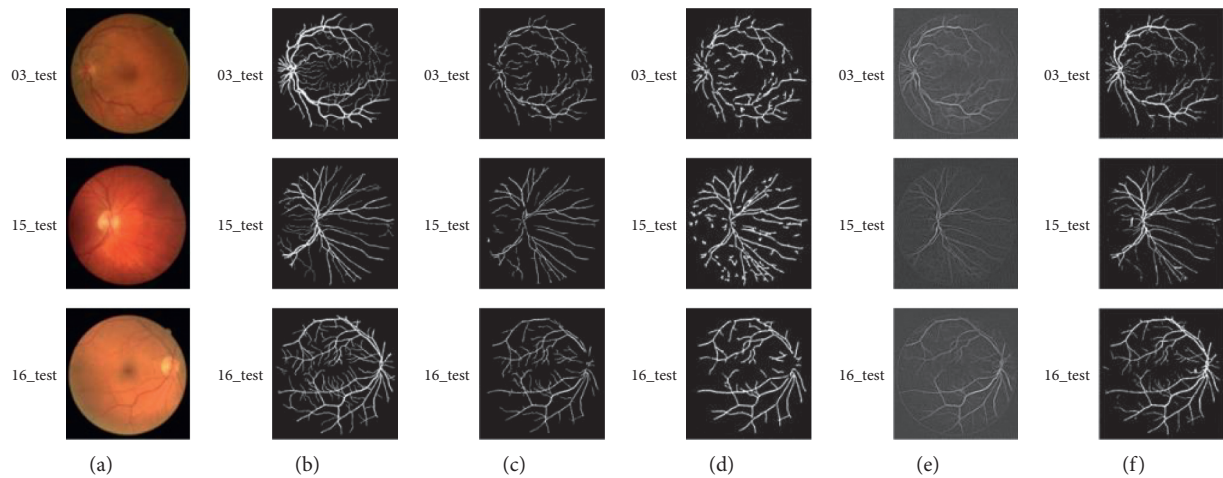


FIGURE 10: Comparison of different segmentation methods for the DRIVE fundus database. (a) Original image, (b) ground truth, (c) Zhang [31], (d) Jiang and Mojon [39], (e) Zana and Klein [6], and (f) our algorithm.

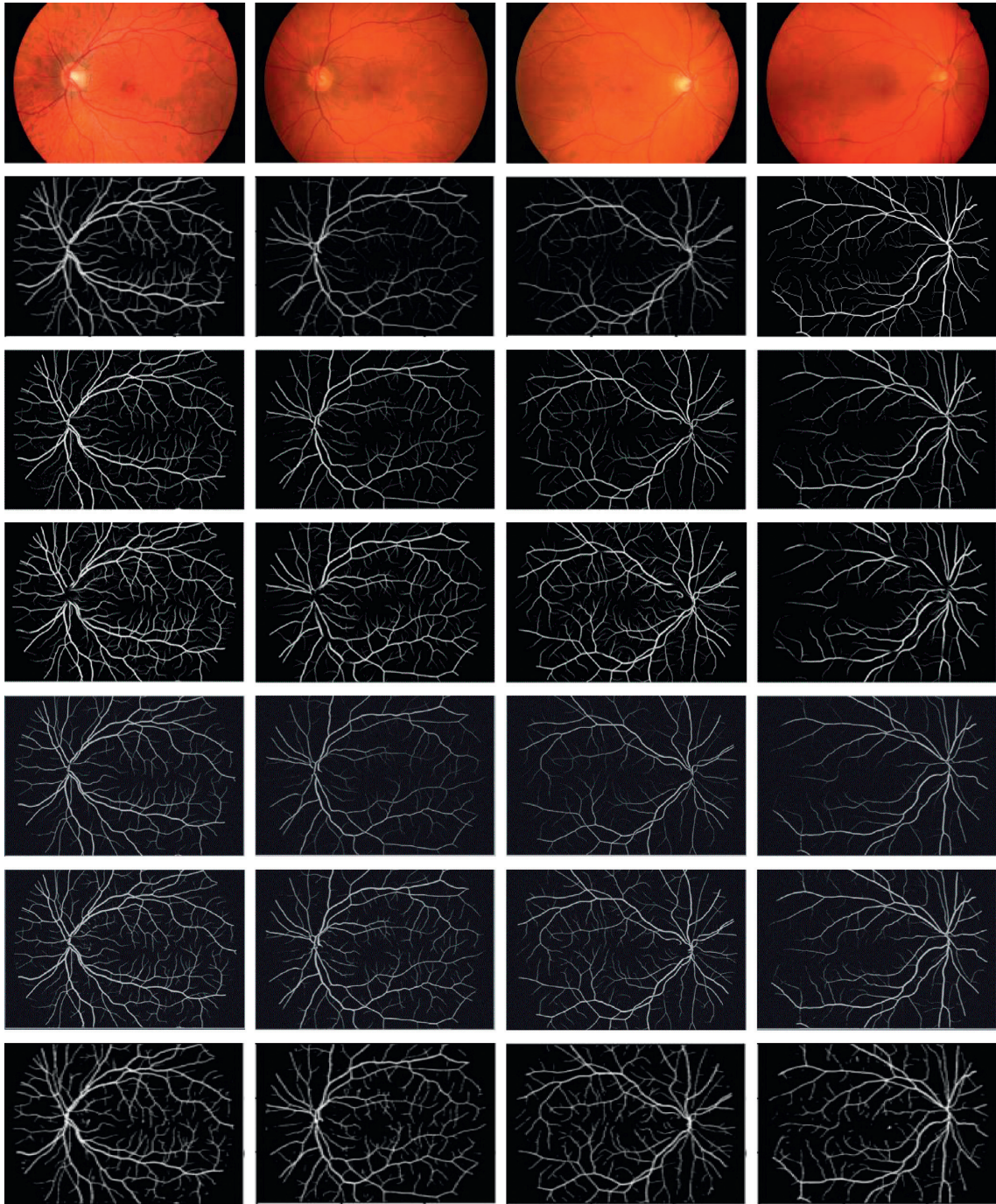


FIGURE 11: Vessel segmentation results on the HRF dataset, from up to down is input images, ground truth, AG-UNet, IterNet, M-GAN, NMF+3D U-Net, and ours, respectively.

TABLE 4: Performance comparison of different methods for the DRIVE fundus database.

Method	Algorithm	ACC	TPR	TNR
Unsupervised method	Vlachos and Dermatas [12]	0.9290	0.7470	—
	Zana and Klein [6]	0.9377	0.6971	—
	Mendonça and Campilho [8]	0.9452	0.7344	0.9764
	Jiang and Mojon [39]	0.9212	0.6399	—
	You et al. [44]	0.9434	0.7410	—
	Chaudhuri et al. [45]	0.8773	0.2716	0.9785
	Zhang et al. [46]	0.9476	0.7743	0.9725
	Zhao et al. [47]	0.9540	0.7420	0.9820

TABLE 4: Continued.

Method	Algorithm	ACC	TPR	TNR
Supervised method	Fraz et al. [5]	0.9480	0.7406	0.9807
	Lin et al. [27]	0.9536	0.7632	—
	Fu et al. [26]	0.9470	0.7294	—
	Niemeijer et al. [7]	0.9416	—	—
	Maji et al. [21]	0.9470	—	—
	Soares et al. [48]	0.9466	0.7332	0.9782
	Staal et al. [49]	0.9442	0.7322	0.9646
	Zhu et al. [50]	0.9607	0.7528	0.9820
	Marin et al. [18]	0.9452	0.7067	0.9801
	GAN [51]	0.9562	0.7746	0.9753
	Wang et al. [20]	0.9767	0.8173	0.9733
Khowaja et al. [52]	0.9753	0.8176	0.9709	
Our algorithm	0.9701	0.8011	0.9849	

TABLE 5: Performance comparison of different methods for the STARE fundus database.

Method	Algorithm	ACC	TPR	TNR
Unsupervised method	Zhang et al. [46]	0.9554	0.7791	0.9758
	Hoover et al. [13]	0.9264	0.6747	0.9565
	Mendonça and Campilho[8]	0.9440	0.6996	0.9730
	Jiang and Mojon [39]	0.9009	—	—
	You et al. [44]	0.9497	0.7260	—
	Zhao et al. [47]	0.9560	0.7800	0.9780
Supervised method	Fraz et al. [5]	0.9534	0.7548	0.9763
	Lin et al. [27]	0.9603	0.7423	—
	Fu et al. [26]	0.9545	0.7140	—
	Niemeijer et al. [7]	0.9534	0.7548	0.9763
	Soares et al. [48]	0.9480	0.7207	0.9747
	Staal et al. [49]	0.9516	—	—
	Marin et al. [18]	0.9526	0.6944	0.9819
	GAN [51]	0.9647	0.7940	0.9869
	Wang et al. [20]	0.9813	0.8104	0.9791
	Khowaja et al. [52]	0.9751	0.8239	0.9749
Our algorithm	0.9683	0.6032	0.9967	

TABLE 6: Performance comparison of different methods for the HRF fundus database.

Method	ACC	TPR	TNR
AG-UNet [40]	0.9600	0.8297	0.9875
IterNet [42]	0.9623	0.7524	0.9743
M-GAN [41]	0.9700	0.6948	0.9931
NMF+3D U-Net [43]	0.9688	0.7451	0.9903
Our algorithm	0.9698	0.7478	0.9895

TABLE 7: Performance analysis of different modified U-Nets on DRIVE and STARE databases with respect to the measuring metrics.

Method	DRIVE			STARE		
	ACC	TPR	TNR	ACC	TPR	TNR
MSFFU-Net [32]	0.9694	0.7762	0.9835	0.9537	0.7721	0.9885
Dense U-net [34]	0.9511	0.7986	0.9736	0.9538	0.7914	0.9722
AA-UNet [53]	0.9558	0.7941	0.9798	0.9640	0.7598	0.9878
DUNet [54]	0.9566	0.7963	0.9800	0.9641	0.7595	0.9878
EEA U-net [55]	0.9577	0.7918	0.9708	0.9445	0.8021	0.9561
Our algorithm	0.9701	0.8011	0.9849	0.9683	0.6032	0.9967

4. Conclusions

The segmentation of retinal vascular images automatically plays a vital role in computer-aided fundus diagnosis and disease screening. There are some differences between retinal vessels and background. Nevertheless, the contrast is not strong enough, which makes the accurate segmentation of retinal vessels harder [56]. Therefore, the segmentation of fundus retinal vessels automatically is a complex and challenging mission. To deal with the problem that the accuracy is not enough in retinal blood vessel segmentation, a new algorithm that could segment retinal blood vessel automatically comes up. The algorithm is based on an improved U-Net segmentation model in this paper. Firstly, the RGB fundus image is preprocessed; secondly, the U-Net model is constructed. The extracted image block could be used during the training stage in the U-Net model; finally, the trained U-Net model is tested to obtain the final results to achieve automatic fundus image segmentation. The experimental results of retinal blood vessel segmentation on the DRIVE, STARE, and HRF color fundus image database show that the algorithm has high accuracy in extracting retinal vessels and can retain more microscopic vessels and complete network. This algorithm can provide necessary theoretical and technical support for ophthalmologists to track the development of fundus lesions and reveal the pathogenesis with a particular clinical reference value.

Although the method of the U-Net network proposed in our paper has made some achievements in the segmentation of retinal vessels, there are still some problems and challenges. First of all, the number of retinal fundus images is limited. Due to ethical issues such as protecting patients' privacy, it is not easy to obtain fundus images, and the amount of data is not abundant. Simultaneously, the ground truth corresponding to RGB fundus images relies on expert observation and manual segmentation, so the database arrangement is tedious. The amount of data in the existing public fundus image database is small, which has some limitations to the training model. Secondly, the RGB fundus image in the image set is affected by uneven illumination and low image resolution, which cannot guarantee the effective recognition of blood vessels. Given the shortcomings, our research's subsequent work is to integrate or postprocess the results to further optimize the segmentation results based on the U-Net network combined with effective technologies and methods.

Data Availability

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

Conflicts of Interest

No conflicts of interest are declared by the authors.

Acknowledgments

This work was funded in part by the National Natural Science Foundation of China (Grant no. 62072413) and also

supported by the Natural Science Foundation of Zhejiang Province of China (Grant no. LY16F010008).

References

- [1] N. I. Kuryshva, "Selective elective ively national natural science foundation of China," *Vestnik Oftalmologii*, vol. 135, no. 2, pp. 144–150, 2019.
- [2] S. S. Kar and S. P. Maity, "Automatic detection of retinal lesions for screening of diabetic retinopathy," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 608–618, 2018.
- [3] C. Bai, L. Huang, X. Pan, J. Zheng, and S. Chen, "Optimization of deep convolutional neural network for large scale image retrieval," *Neurocomputing*, vol. 303, no. 16, pp. 60–67, 2018.
- [4] P. Zhou, C. Bai, J. Xia, and S. Chen, "CMRDF: a real-time food alerting system based on multimodal data," *IEEE Internet of Things Journal*, vol. 99, p. 1, 2020.
- [5] M. M. Fraz, P. Remagnino, A. Hoppe et al., "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, 2012.
- [6] F. Zana and J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1010–1019, 2001.
- [7] M. Niemeijer, J. Staal, B. van Ginneken et al., "Comparative study of retinal vessel segmentation methods on a new publicly available database," *Proceedings of Spie the International Society for Optical Engineering*, vol. 5370, pp. 648–656, 2004.
- [8] A. M. Mendonça and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE Transactions on Medical Imaging*, vol. 25, no. 9, pp. 1200–1213, 2006.
- [9] S. Garg, S. Jayanthi, and S. Chandra, "Unsupervised curvature-based retinal vessel segmentation," in *Proceedings of the International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 344–347, IEEE, Washington, DC, USA, April 2007.
- [10] O. Chutatape, Z. Liu, and S. M. Krishnan, "Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters," in *Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society*, vol. 6, pp. 3144–3149, Hong Kong Sar, China, November 1998.
- [11] L. Gang, O. Chutatape, and S. M. Krishnan, "Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 2, pp. 168–172, 2002.
- [12] M. Vlachos and E. Dermatas, "Multi-scale retinal vessel segmentation using line tracking," *Computerized Medical Imaging and Graphics*, vol. 34, no. 3, pp. 213–227, 2010.
- [13] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, pp. 203–210, 2000.
- [14] M. Zardadi, N. Mehrshad, and S. M. Razavi, "Unsupervised segmentation of retinal blood vessels using the human visual system line detection model," *Information Systems & Telecommunication*, vol. 4, pp. 125–133, 2016.
- [15] C. L. Srinidhi, P. Aparna, and J. Rajan, "Automated method for retinal artery/vein separation via graph search

- metaheuristic approach,” *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 2705–2718, 2019.
- [16] K. Upadhyay, M. Agrawal, and P. Vashist, “Unsupervised multiscale retinal blood vessel segmentation using fundus images,” *IET Image Processing*, vol. 14, no. 11, pp. 2616–2625, 2020.
- [17] S. Aslani and H. Sarnel, “A new supervised retinal vessel segmentation method based on robust hybrid features,” *Biomedical Signal Processing and Control*, vol. 30, pp. 1–12, 2016.
- [18] O. C. Marin, M. G. Penedo, M. Penas, M. J. Carreira, and F. Gonzalez, “Personal authentication using digital retinal images,” *Pattern Analysis and Applications*, vol. 9, no. 1, pp. 21–33, 2006.
- [19] Y. Jiang, N. Tan, T. Peng, and H. Zhang, “Retinal vessels segmentation based on dilated multi-scale convolutional neural network,” *IEEE Access*, vol. 7, pp. 76342–76352, 2019.
- [20] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, “Hierarchical retinal blood vessel segmentation based on feature and ensemble learning,” *Neurocomputing*, vol. 149, pp. 708–717, 2015.
- [21] D. Maji, A. Santara, P. Mitra et al., “Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images in Engineering in Medicine and Biology Society (EMBC),” in *Proceedings of the 2015 37th Annual International Conference of the IEEE*, pp. 3029–3032, Milan, Italy, August 2015.
- [22] A. Che, M. Zulfaezal, A. Hamid et al., “Supervised retinal vessel segmentation based on neural network using broader aging dataset,” in *Proceedings of the IWBBIO*, pp. 1235–1242, Granada, Spain, April 2014.
- [23] P. Liskowski and K. Krawiec, “Segmenting retinal blood vessels with new deep neural networks,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 11, pp. 2369–2380, 2016.
- [24] J. H. Tan, H. Fujita, S. Sivaprasad et al., “Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network,” *Information Sciences*, vol. 420, pp. 66–76, 2017.
- [25] Y. Wu, Y. Xia, Y. Song et al., “Multiscale network followed network model for retinal vessel segmentation,” *Medical Image Computing and Computer Assisted Intervention-MICCAI 2018*, Springer, Cham, Switzerland, 2018.
- [26] H. Fu, Y. Xu, D. W. K. Wong et al., “Retinal vessel segmentation via deep learning network and fully-connected conditional random fields,” in *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 698–701, Washington, DC, USA, April 2016.
- [27] Y. Lin, H. Zhang, and G. Hu, “Automatic retinal vessel segmentation via deeply supervised and smoothly regularized network,” *IEEE Access*, vol. 7, pp. 57717–57724, 2018.
- [28] T. M. Khan, M. Alhusein, K. Aurangzeb, M. Arsalan, S. S. Naqvi, and S. J. Nawaz, “Residual connection-based encoder decoder network (RCED-Net) for retinal vessel segmentation,” *IEEE Access*, vol. 8, pp. 131257–131272, 2020.
- [29] C. Wu, Y. Zou, and Z. Yang, “U-GAN: generative adversarial networks with U-net for retinal vessel segmentation,” in *Proceedings of the IEEE International Conference on Computer Science & Education (ICCSE)*, pp. 642–646, Jinan, China, 2019.
- [30] T. Laibacher, T. Weyde, and S. Jalali, “M2U-Net: effective and efficient retinal vessel segmentation for real-world applications,” in *Proceedings of the CVPR Workshops*, Long Beach, CA, USA, 2019.
- [31] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, “UNet++: redesigning skip connections to exploit multiscale features in image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 1856–1867, 2020.
- [32] D. Yang, G. Liu, M. Ren, B. Xu, and J. Wang, “A multi-scale feature fusion method based on U-net for retinal vessel segmentation,” *Entropy*, vol. 22, no. 8, p. 811, 2020.
- [33] L. Yicceng, X. Rui, Z. Li et al., “Feature pyramid U-net for retinal vessel segmentation,” *IET Image Processing*, vol. 22, no. 2, 2021.
- [34] C. Wang, Z. Zhao, Q. Ren, Y. Xu, and Y. Yu, “Dense U-net based on patch-based learning for retinal vessel segmentation,” *Entropy*, vol. 21, no. 2, p. 168, 2019.
- [35] J. Hu, H. Wang, S. Gao et al., “S-UNet: a bridge-style U-net framework with a saliency mechanism for retinal vessel segmentation,” *IEEE Access*, vol. 99, p. 1, 2019.
- [36] G. A. Francia, C. Pedraza, M. Aceves et al., “Chaining a U-net with a residual U-net for retinal blood vessels segmentation,” *IEEE Access*, vol. 8, pp. 38493–38500, 2020.
- [37] O. Ronneberger, P. Fischer, and T. Brox, “U-net: convolutional networks for biomedical image segmentation,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, Munich, Germany, October 2015, Lecture Notes in Computer Science.
- [38] M. Z. Alom, M. Hasan, C. Yakopcic et al., “Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation,” 2018, <https://arxiv.org/abs/1802.06955>.
- [39] X. Jiang and D. Mojon, “Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images,” *Pattern Analysis & Machine Intelligence IEEE Transactions on*, vol. 25, no. 1, pp. 131–137, 2003.
- [40] J. Schlemper, O. Oktay, M. Schaap et al., “Attention gated networks: learning to leverage salient regions in medical images,” *Medical Image Analysis*, vol. 53, pp. 197–207, 2019.
- [41] K.-B. Park, S. H. Choi, and J. Y. Lee, “M-GAN: retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks,” *IEEE Access*, vol. 8, pp. 146308–146322, 2020.
- [42] L. Li, M. Verma, Y. Nakashima et al., “IterNet: retinal image segmentation utilizing structural redundancy in vessel networks,” in *Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 3645–3654, Snowmass Village, CO, USA, March 2020.
- [43] Y. Yu and H. Zhu, “Retinal vessel segmentation with constrained-based nonnegative matrix factorization and 3D modified attention U-Net,” *EURASIP Journal on Image and Video Processing*, vol. 2021, no. 1, pp. 1–21, 2021.
- [44] X. You, Q. Peng, Y. Yuan, Y.-m. Cheung, and J. Lei, “Segmentation of retinal blood vessels using the radial projection and semi-supervised approach,” *Pattern Recognition*, vol. 44, no. 10–11, pp. 2314–2324, 2011.
- [45] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, “Detection of blood vessels in retinal images using two-dimensional matched filters,” *IEEE Transactions on Medical Imaging*, vol. 8, no. 3, pp. 263–269, 1989.
- [46] J. Zhang, E. Bekkers, S. Abbasi et al., “Robust and fast vessel segmentation via Gaussian derivatives in orientation scores,” in *Proceedings of the International Conference on Image Analysis and Processing*, Niagara Falls, Canada, May 2015.
- [47] Y. Zhao, L. Rada, K. Chen, S. P. Harding, and Y. Zheng, “Automated vessel segmentation using infinite perimeter

- active contour model with hybrid region information with application to retinal images,” *IEEE Transactions on Medical Imaging*, vol. 34, no. 9, pp. 1797–1807, 2015.
- [48] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, “Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification,” *IEEE Transactions on Medical Imaging*, vol. 25, no. 9, pp. 1214–1222, 2006.
- [49] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, “Ridge-based vessel segmentation in color images of the retina,” *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004.
- [50] C. Zhu, Y. Xiang, B. Zou et al., “Retinal vessel segmentation in fundus images using CART and AdaBoost,” *Journal of Computer-Aided Design and Computer Graphics*, vol. 26, no. 3, pp. 445–451, 2014.
- [51] W. Abbas, M. H. Shakeel, N. Khurshid et al., “Patch-based generative adversarial network towards retinal vessel segmentation,” in *Proceedings of the Neural Information Processing*, Vancouver, Canada, December 2019.
- [52] S. A. Khowaja, P. Khuwaja, and I. A. Ismaili, “A framework for retinal vessel segmentation from fundus images using hybrid feature set and hierarchical classification,” *Signal, Image and Video Processing*, vol. 13, no. 2, pp. 379–387, 2019.
- [53] Y. Lv, H. Ma, J. Li, and S. Liu, “Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation,” *IEEE Access*, vol. 8, pp. 32826–32839, 2020.
- [54] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, “DUNet: a deformable network for retinal vessel segmentation,” *Knowledge-Based Systems*, vol. 178, no. 15, pp. 149–162, 2019.
- [55] V. Sathananthavathi and G. Indumathi, “Encoder enhanced atrous (EEA) unet architecture for retinal blood vessel segmentation,” *Cognitive Systems Research*, vol. 67, no. 4, pp. 84–95, 2021.
- [56] Z. Gao, Y. Li, Y. Sun et al., “Motion tracking of the carotid artery wall from ultrasound image sequences: a nonlinear state-space approach,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 1, pp. 273–283, 2017.

Research Article

Application of Internet of Things Compressed Sensing and Information Interaction Technology in Intelligent Transportation Layout

Na Li , Ze Wu , and Zhongbiao Zhao 

School of Electrical and Mechanical Engineering, Xuchang University, Xuchang 461000, China

Correspondence should be addressed to Na Li; ln@xcu.edu.cn

Received 8 March 2021; Revised 7 April 2021; Accepted 14 April 2021; Published 23 April 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Na Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous development of economy, the number of motor vehicles has increased sharply, and urban traffic congestion has become more and more frequent. Urban traffic congestion has become one of the important reasons that hinder the continuous development of major cities. In order to better solve the problem of urban traffic congestion, it is more urgent to build urban intelligent transportation system. The technology related to Internet of things is becoming more and more mature, which has become a new idea of collecting traffic information, providing convenience for traffic, and injecting vitality into the study of intelligent transportation layout. The traffic system in the city or the main traffic intersection, through the traffic guidance layout road, provides the traffic instruction for the passenger on the highway network, enables the driver to choose the suitable driving road, can provide the travel guidance service for the driver, adjusts the flow distribution, and improves the traffic condition. This paper describes the research status of intelligent transportation layout and uses the basic theory of compressed sensing and information interaction technology to carry out data fusion and reconstruction algorithm in intelligent transportation. Finally, the application of compressed sensing in intelligent transportation layout is summarized to realize data collection in intelligent transportation network. In this paper, compressed sensing theory is applied to the data acquisition of intelligent transportation network to reduce the amount of data and improve the effective data acquisition. The reconstruction algorithm is used to realize the data reconstruction to ensure the accuracy and stability of signal reconstruction and the accuracy of network transmission information.

1. Introduction

With the rapid development of domestic economy and the continuous improvement of people's income, China's urbanization level has reached an unprecedented height. With the rapid expansion of urban area and the rising of housing prices, urban population is constantly moving to the suburbs. In order to improve the efficiency of travel and shorten the consumption of time on the way to work, more and more families use cars as a necessary means of travel, resulting in a sharp increase in the number of urban motor vehicles [1]. Although the number of urban traffic roads has increased and the carrying capacity of existing urban traffic roads has also been improved, it still cannot meet the demand for motor vehicles for roads in the same period. Therefore,

urban traffic roads are becoming more and more crowded. The phenomenon of urban traffic congestion has attracted the attention of governments all over the world and has become an urgent problem for traffic managers in major cities [2]. Urban traffic congestion has a serious impact on the daily travel of urban residents and has become one of the main factors hindering the further improvement of the level of urbanization. Traffic congestion not only reduces the traffic capacity of the road but also shortens the service life of the road and greatly increases the traffic delay time of urban residents [3]. In addition, urban traffic congestion will lead to a series of problems, which has brought great threat to the life and property of the country and the people and has become the focus of public opinion. Aiming at a series of major problems caused by urban traffic congestion,

governments around the world deepen cooperation in this field [4]. At the same time, a large number of researchers have also invested in the research of urban traffic congestion. Intelligent transportation layout has greatly expanded people's traditional understanding of urban transportation and opened up new ideas for the future planning of urban transportation. With the development of its research, it has been widely considered as one of the best ways to solve the problem of urban traffic congestion. In order to solve the problem of urban traffic congestion, the governments of various countries also continue to increase the investment in infrastructure construction, so as to prepare for the comprehensive completion of the intelligent transportation system [5].

The construction of intelligent transportation system has become the only way for the sustainable development of urban transportation. Intelligent transportation will be the future development direction of transportation. It will make the traditional transportation mode more intelligent, safer, energy-saving, and efficient through high technology. Internet of things is an important part of the new generation of information technology; through radio frequency identification, global positioning system, and other types of information sensing equipment, according to the agreed protocol, any object is connected to the Internet for information exchange and communication. With the continuous development of Internet of things technology, it also injects new power into the further development and improvement of its. In view of the serious harm of urban traffic congestion to urban development, this paper proposes an intelligent transportation layout based on Internet of things compressed sensing and information interaction technology to realize the traffic layout estimation method based on compressed sensing. Experimental results show that the sparse coding method can achieve more effective sparse representation of traffic layout information at adjacent times and can obtain lower estimation error. In order to further reduce the estimation error of compressed sensing method for urban traffic layout, this paper proposes a sparse representation matrix based on Gaussian kernel function by combining the road network topology and introducing the Gaussian joint distribution model, so as to improve the sparsity performance of traffic layout information, and proposes a method to estimate the traffic layout by using part of the detection vehicles; this method constructs a sparse representation matrix; it analyses whether the sparse representation matrix and measurement matrix constructed by this method meet the dual incoherence principle in compressed sensing theory. The simulation results of real data show that this method can reduce the reconstruction error of traffic layout information and achieve more accurate estimation of urban traffic layout.

This paper mainly describes the content of the paper from six parts, and the specific arrangement of each chapter is summarized as follows: the first section is the introduction, which mainly introduces the specific topic background and research significance of the paper,

analyses the research status of compressed sensing theory and information interaction technology in the application of Internet of things at home and abroad, and summarizes the main research content and paper structure arrangement of this paper. The second section discusses the related work. The third section introduces the Internet of things compressed sensing model and information interaction technology and analyses and verifies the usability, efficiency, flexibility, and security of the proposed model from the theoretical and experimental aspects. The fourth section discusses the application of compressed sensing and information interaction technology based on the Internet of things in intelligent transportation, from the aspects of compressed sensing data collection and application effectiveness. In Section 5, the performance of this method is verified by simulation. The sixth section is the summary and prospect of this paper.

2. Related Work

As an important basis for the implementation of traffic road planning, traffic layout makes a large number of scholars invest in the study of traffic layout estimation methods [6]. The main research methods can be divided into three categories: one is based on data fusion; the other is based on traffic information matching and tracking; the third is based on traffic dynamics. The technology related to Internet of things is becoming more and more mature, which has become a new idea of collecting traffic information, providing convenience for traffic, and injecting vitality into the study of intelligent transportation layout [7]. Intelligent transportation will be the future development direction of transportation. It will make the traditional transportation mode more intelligent, safer, energy-saving, and efficient through high technology [8]. In recent years, the application of compressed sensing in intelligent transportation layout has received extensive attention and development.

A large number of data simulation experiments have been carried out in the research of road traffic layout estimation methods, and the advantages and disadvantages of multiple regression, fuzzy reasoning, neural network, and speed integral on the traffic layout estimation model are analysed [9]. Some researches propose data fusion traffic estimation method based on robust Kalman filter algorithm. This method mainly uses the existing road GPS data information to estimate the road traffic layout with missing GPS data, which improves the integrity of road traffic layout estimation. Some scholars use Bayesian estimation, fuzzy logic, and other data fusion methods to establish a multiheterogeneous traffic information three-level fusion system and obtain high-precision and high-reliability traffic layout data [10]. Other studies have proposed real-time traffic layout estimation method based on floating vehicle data. Its main work is map matching processing, road section division method, and minimum sample size research on GPS data of service vehicles. In this paper, a series of data sources such as special

vehicle detector, vehicle tracking and positioning system, and road toll system are studied, and the traffic data preprocessing and multidata fusion traffic layout monitoring method using multineural network are proposed [11].

According to the method of estimating road flow and density with detection data, through a large number of simulation studies, when the proportion of detection vehicles is 3.5% and 0.2%, this method can accurately output traffic flow of 5 minutes and hours. In part of the study, we proposed a method to estimate the traffic layout of different roads by accurately estimating the average speed of road sections and obtained the estimation accuracy of 3% floating vehicles in congested flow which is equivalent to that of 30% floating vehicles in free flow through experiments. It is mentioned in some literature that the travel time of all road sections can be estimated by the historical data of road traffic, and the weight parameter mechanism is proposed to overcome the data sparsity [12]. In this paper, the idea of using probe vehicle to estimate urban traffic layout by using compressed sensing method is proposed. The method of traffic layout estimation based on the combination of probe vehicle and compressed sensing can reduce the cost of data collection [13]. With the in-depth study of traffic layout estimation methods by a large number of scholars, more and more traffic layout estimation methods have been proposed. Especially in the face of such a large amount of traffic flow data, compressed sensing method can greatly reduce the amount of data required for traffic layout estimation and can solve the problem of missing road data, which provides great technical support for the establishment of urban intelligent transportation system.

The application of compressed sensing theory has attracted the attention of many scholars in the field of medical imaging, astronomical observation, and speech compression. Aiming at the problem of a large amount of data information in wireless sensor network, some scholars use compressed sensing technology to observe the data of a single sensor node and carry out fusion processing in the process of data transmission, so as to realize the application of compressed sensing in wireless sensor network data collection [14, 15]. Their research also shows that its application can significantly reduce the data storage and data transmission time; domestic scholars have also carried out research on compressed sensing theory. In the field of sensor networks, some people have studied efficient data transmission schemes in wireless sensor networks, others have applied compressed sensing to target positioning in wireless sensor networks, and others have applied compressed sensing to reduce low power consumption in body area networks. According to the literature review, there are no reports about compressed sensing to deal with the interactive sensing of massive vehicle sensing information. Many famous universities, including MIT, Harvard University, Carnegie Mellon University, Chinese Academy of Sciences, and Tsinghua University, have set up research groups on the theory and application of compressed sensing. CS has rapidly become a hot research field in recent years, and related research results emerge endlessly, which has been widely used in cognitive radio medical imaging, signal

processing, and target setting in many fields, such as bit, wireless communication, radar detection, channel coding, channel estimation, geological exploration, and satellite remote sensing image [16]. In the current situation, although the expansion of roads and the control of the increase in the number of vehicles can alleviate the traffic pressure, the utilization of land resources is limited after all. This way cannot solve the congestion phenomenon from the perspective of sustainable development. This paper describes the research status of intelligent transportation layout and uses the basic theory of compressed sensing and information interaction technology to carry out data fusion and reconstruction algorithm in intelligent transportation.

3. Internet of Things Compressed Sensing and Information Interaction Technology

3.1. Compressed Sensing of Internet of Things. If the coefficient of the signal x in the basis vector has the following properties, it is said that the signal x is compressible in the basis vector: the absolute values of the coefficients of W are arranged from large to small. Generally, if a signal contains redundancy, it is compressible. Therefore, for a node data field related to data, we can reasonably assume that the signal is compressible. The n -dimensional vector represents the approximation matrix of the largest m elements, the remaining elements are set to zero and similarly introduced to represent the approximation matrix of the smallest elements, and the remaining elements are set to zero.

$$G = \varphi \varepsilon_s + \varphi \varepsilon_e. \quad (1)$$

It is obtained by random sampling. We assume that the compressed sensing data is noise data, which is represented by N and obeys the Gaussian distribution of zero means.

$$G = \varphi \varepsilon_s + n, \quad (2)$$

where n is a Gaussian distribution with zero mean and σ^2 variance. So, we get Gaussian likelihood estimation.

$$H(G|\varepsilon_s, \sigma^2) = \frac{e^{-\|G - \varphi \varepsilon_s\|^2 / \sigma^2}}{\pi \sigma^2}. \quad (3)$$

Considering the accuracy of the Gaussian density function, we can get the following results:

$$H(\varphi|\alpha) = \prod_{i=1}^T T(\varphi_i|0, \alpha_i) + \Gamma(\alpha_i|\tau, \nu). \quad (4)$$

Ignoring the influence of α , the weight can be calculated by the following formula:

$$H(\varphi|\tau, \nu) = \prod_{i=1}^T \int_0^\infty T(\varphi_i|0, \alpha_i) + \Gamma(\alpha_i|\tau, \nu) d\alpha_i. \quad (5)$$

Through the above analysis, the research of coefficient degree in compressed sensing is transformed into a linear regression problem. The prior constraint of this problem is sparse. The widely used sparse promotion algorithm is the Laplace density function.

$$H(\varphi|\lambda) = \prod_i \int H(\varphi_i|\gamma_i)H(\gamma_i|\lambda)d\gamma_i. \quad (6)$$

Compressed sensing can be used to construct classifiers. According to the theory of compressed sensing, the signal can be compressed, or the signal is sparse in a certain sensing field. The sparse signal can be projected into a low dimensional space. Through optimization, the high probability signal in the projection can be reconstructed into the original signal. The reconstruction signal contains a lot of information and effectively distinguishes the original signal; the principle of its main implementation is shown in Figure 1.

In signal compression, the signal is first transformed, such as discrete cosine transform or wavelet transform, then a few coefficients with larger absolute value are compressed and coded, and the zero or near zero coefficients are discarded [17]. The measured value is not the signal itself, but the projection value from high dimension to low dimension. From the mathematical point of view, each measured value is the combination function of each sample signal in the traditional sense; that is, a measured value contains a small amount of information of all sample signals [18]. By compressing the data, most of the sampled data is discarded, but the perception effect is not affected. For example, when a digital camera with millions of pixels is used to image the scene, a large amount of pixel information will be obtained, but only a small amount of information will be stored and transmitted after compression coding, and finally, the original image will be reconstructed through the corresponding decompression algorithm. If the signal itself is compressible, can we directly obtain its compressed representation (i.e., compressed data) so as to omit the sampling of a large amount of redundant information? Cand proved in 2006 that the original signal can be accurately reconstructed from some Fourier transform coefficients, which laid a theoretical foundation for compressed sensing. Then Cand *é* s and Donohoe put forward the concept of compressed sensing on the basis of related research. The core idea is that sampling and compression coding occur in the same step. Firstly, the nonadaptive linear projection of the signal is collected, and then the original signal is reconstructed from the measured value according to the corresponding reconstruction algorithm. Modules of information interaction system were shown in Figure 2.

Through research, it has been known that as long as the signal is compressible or sparse in a certain transform domain, then the high-dimensional signal obtained from the transformation can be projected onto a low dimensional space by using an observation matrix not related to the transform base, and then the original signal can be reconstructed with high accuracy from these few projections by solving an optimization problem, which can prove such a projection value package. It contains enough information of reconstructed signal; in this theoretical framework, the sampling rate is mainly determined by the sparsity and noncorrelation of information in the signal, not by the bandwidth of the signal.

3.2. Information Interaction Technology. Information exchange of Internet of things is a process of information transmission, information sharing, and information

exchange based on network system with many heterogeneous network nodes [19]. Through information exchange, each node of Internet of things can obtain the information of environment and other nodes intelligently and autonomously. Although the existing research work has carried out in-depth research on human-computer interaction theory of traditional information system and put forward a complete information exchange model interaction model, there is no mature theoretical system for information interaction of Internet of things [20]. Through in-depth study of human-computer interaction process of large-scale information system, the model thinks that information interaction is completed by the interaction among three basic objects: user, system, and content. The fundamental purpose of user using information system is to use the content of the system, but the user should obtain the information successfully content and must use the function provided by the system to complete the corresponding system operation, and the content is the system as a carrier for different applications to display information. Based on the above information interaction model, combined with the characteristics of the Internet of things, this paper puts forward the information interaction model of the Internet of things as shown in Figure 3; the basic object of the model is composed of three parts: user, network, and content different from the meaning of user in the traditional information interaction model; the user here is a generalized user, including not only the traditional human-computer interaction user but also the sink node, cluster head node, routing node, and general network node. The information interaction system of the Internet of things refers to the perception network itself; that is, the whole system including information perception unit, operation and storage unit, and energy unit information space based on the Internet of things network system constitutes the content of information interaction, including various perception data of network nodes, network state information, and high-level semantic and event information that users are interested in sink node obtains the interaction process of node perception information through the perception network [21].

4. Application of Compressed Sensing and Information Interaction Technology Based on Internet of Things in Intelligent Transportation

4.1. Acquisition of Compressed Sensing Data. In the vehicle road coordination system, the vehicle's microprocessor is mainly used to obtain on-board information and analyse and process the underlying information required by various applications in the vehicle road coordination system and put in various intelligent algorithms to comprehensively process all kinds of information, and send the corresponding information to the surrounding vehicles through the vehicle self-organizing network according to the needs. The collection of various heterogeneous sensors on the vehicle makes the vehicle have the ability to collect and identify environmental information, such as the condition

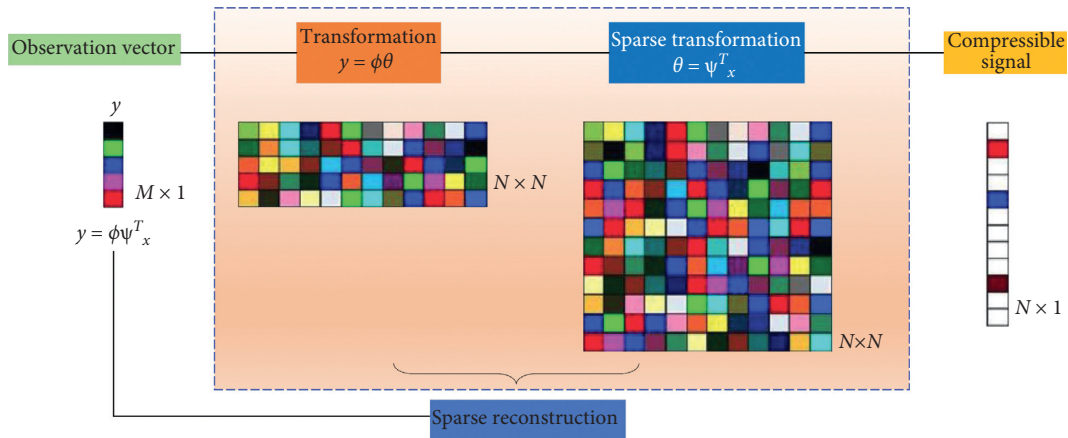


FIGURE 1: Schematic diagram of compressed sensing principle.

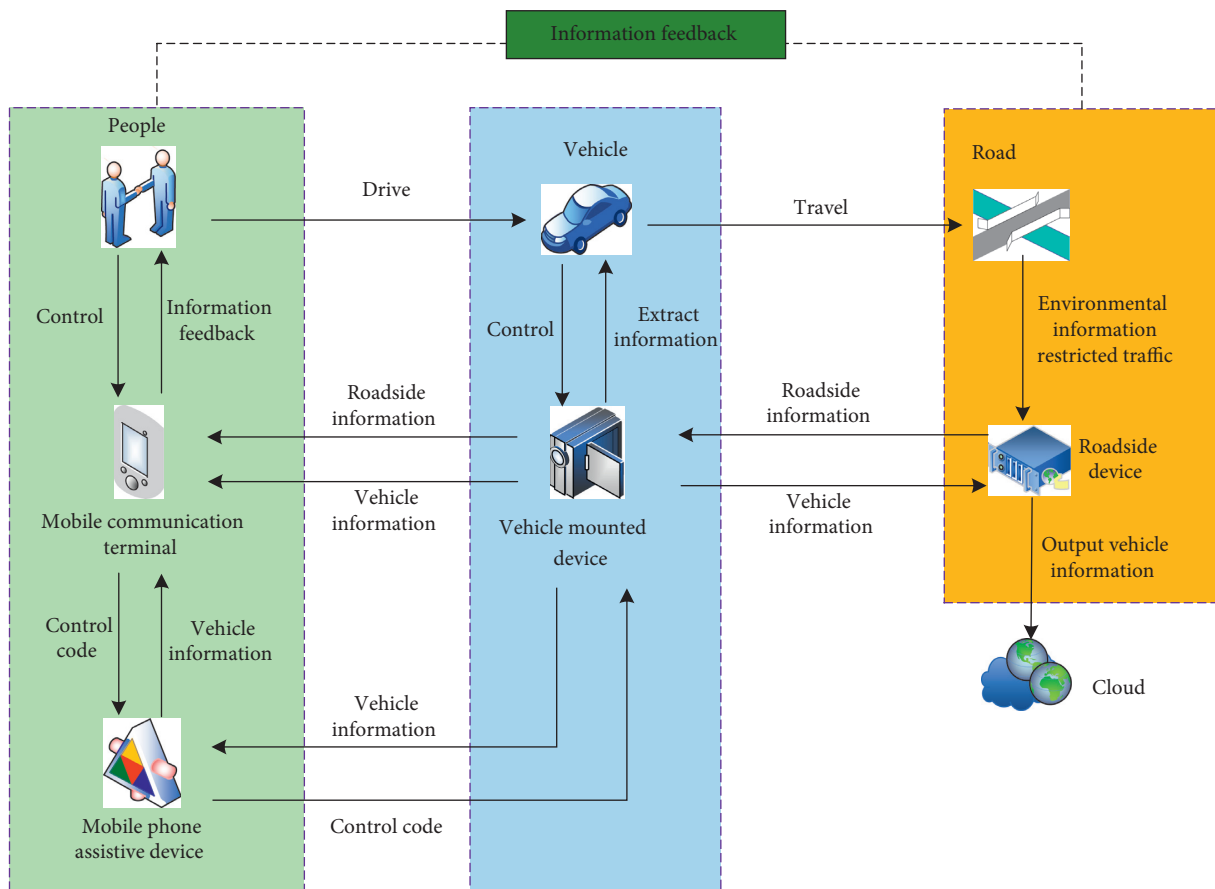


FIGURE 2: Modules of information interaction system.

acquisition of the vehicle running process [22]. With the use of GPS and the installation of acceleration sensor, gyroscope, steering wheel angle sensor, and other types of sensing equipment, the high-precision vehicle motion state information can be obtained. The measurement of vehicle relative positioning information is also detected by the on-board sensors, and distance sensors and vision sensors are mainly used to measure the relative positioning information of the vehicle [23]. The information of the driving environment around the vehicle is mainly obtained by radar, CCD vision

sensor, infrared video sensor, and other types of equipment. Therefore, the ability of vehicle information acquisition under the vehicle road coordination system is greatly increased, which expands the original collection mode and collection range of traffic information. At the same time, the amount of data to be processed is also increasing at an amazing speed, which increases the information processing of the vehicle sensor network and the transmission burden of the vehicle sensor node [24]. As a vehicle sensor node, how to get the least required information in time and reduce

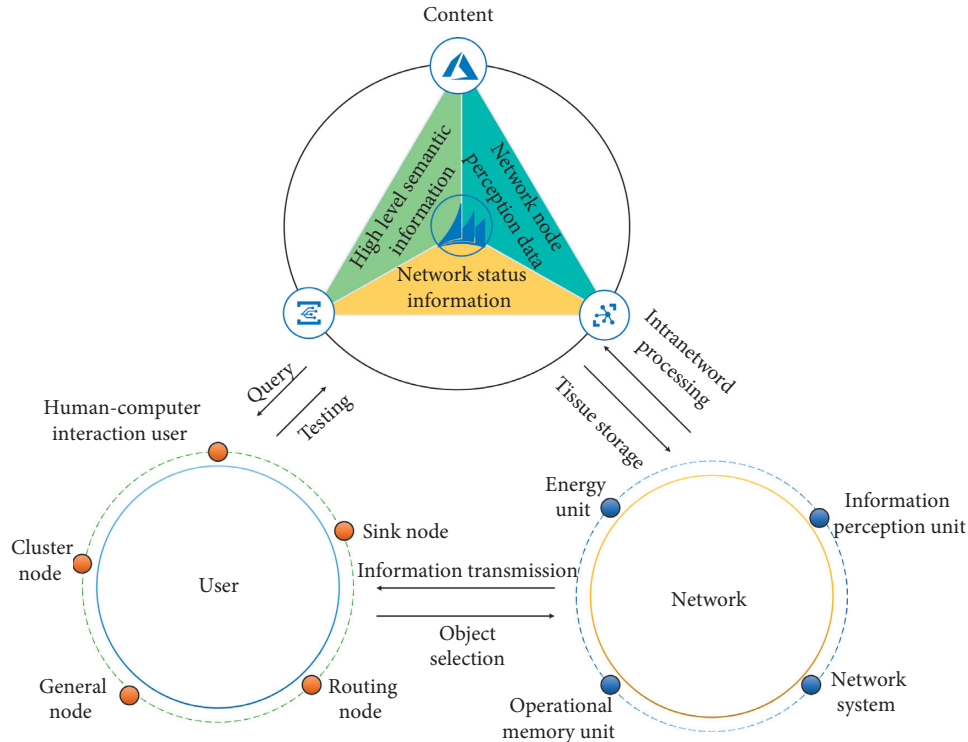


FIGURE 3: Information interaction model of Internet of things.

the cost in order to reduce the waste of broadband caused by unnecessary information collection and transmission, this paper proposes a data compression sensing acquisition method for vehicle sensor network.

The sparse representation of signal is to seek as few basic functions as possible in the transform domain. Fourier transform and wavelet transform are typical signal basis representation methods. The basic idea of wavelet transform is to use a set of wavelets or basis functions to represent a function or signal through scaling and translation. When the signal is represented by wavelet basis, the statistical characteristics of the signal can still be maintained under different scales or compression ratios [25]. Therefore, the original signal detected by vehicle sensor node can be transformed into wavelet domain to meet the prior conditions of compressed sensing, which is also a feasible method to reduce the redundancy of vehicle sensor information. After wavelet transform, the low-frequency part represents the outline of the signal, and the high-frequency part represents the details of the signal [26]. After decomposition, the wavelet coefficients are sorted according to the absolute value, and a certain proportion of the coefficients with relatively large absolute value are retained to obtain the sparse representation of the signal [27]. According to the analysis of vehicle cooperation method, vehicle acceleration information and speed information are important parameters to control vehicle cooperative driving. Figure 4 shows 300 high-frequency sampling points within 5 minutes of vehicle driving speed. It can be seen from Figure 4 that the condition of strict sparse compressed sensing is not satisfied. Compared with the standard Fourier transform, the wavelet

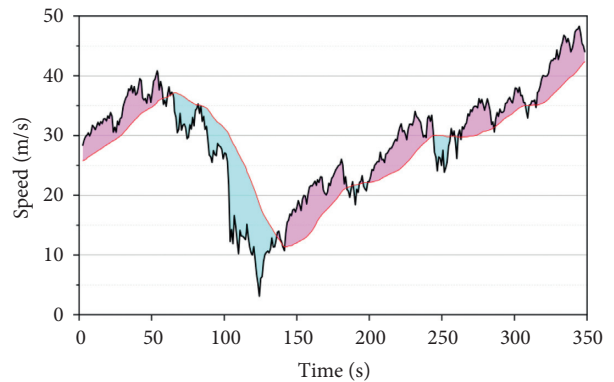


FIGURE 4: Curve of change value of vehicle sensing speed.

functions used in wavelet transform are not unique and diverse. Daubechies function in wavelet theory is a wavelet function constructed by Ingrid Daubechies, a famous wavelet analyst in the world. Except Haar wavelet, other wavelets have no explicit expression, but the square modulus of the transform function is very clear.

For the speed perception of vehicle, the wavelet transform method is used to transform the speed signal, and the corresponding orthogonal transform matrix is obtained. The unified concept of multiresolution analysis in wavelet theory establishes the construction methods of various wavelet bases. Through this analysis method, the orthogonal basis can be constructed for the speed signal perceived by the vehicle. Because of the FIR filter bank corresponding to the tightly supported wavelet base, the construction of wavelet

base is easier to realize. A vehicle sensor signal is analysed by multiresolution method. The scale function K is selected and u is established as a set of orthogonal scale functions. Through the above steps, the wavelet orthogonal basis vector g of vehicle speed information is obtained, which satisfies the sparse condition of compressed sensing.

It can be seen from Figure 5 that the speed values of some continuous high-frequency sampling points are the same, and most of the change values of the vehicle sensing acceleration value should be 0. The high-frequency sampling of vehicle sensing acceleration is shown in Figure 6. Among the 300 sampling points, 100 nonzero points are randomly distributed in the 300 sampling points. The vehicle sensing acceleration signal is strictly sparse, which fully adapts to the compressed sensing theory, so the sparse basis is the unitary matrix.

4.2. Analysis of Application Effectiveness of Compressed Sensing. In order to verify the application effectiveness of compressed sensing theory of vehicle sensing information, MATLAB is used to simulate the compressed sensing of vehicle sensing information. The MSE of the reconstructed results of vehicle sensing speed and acceleration under different measurement times m is displayed. Due to the random uncertainty of Gaussian matrix, the mean square error is obtained by averaging the error obtained from 10 times of simulation for each observation.

For the vehicle sensing velocity and acceleration information in a certain period of time, the observation value m is larger; that is, the more high-frequency sampling points of vehicle sensing data are, the more accurate reconstruction ratio is obtained by simulation. When the m value of random observation is greater than 150, the reconstructed information error is stable. At this time, 150 sampling points were randomly selected from 300 high-frequency sampling points in 5 minutes. The result of information reconstruction is shown in Figure 6. The absolute error of the reconstructed value of a single sampling point can be obtained. The compressed sensing method can use a small amount of data to get more accurate reconstruction and recovery. The reconstruction and recovery error are small, and some error values are close to 0. It is feasible to apply this method to the collection of vehicles sensing data.

5. Analysis of Simulation Experiment

5.1. Validation Analysis under Different Vehicle Density Scenarios. The sparsity of vehicle nodes in its layout is directly related to the stability of its layout state, the connectivity of its routing, and the competitiveness of its channel access. In order to explore the impact of different access protocols and different vehicle density on its layout performance, two scenarios with different vehicle density are designed in this simulation experiment. According to the design of the simulation experiment, the message delay and throughput are simulated. Through the simulation results, it can be concluded that, under the IOT compressed sensing and information interaction technology, the performance of

message delay and throughput is better than that of traditional technology. When the simulation system has 100 vehicle nodes and 50 vehicle nodes, in terms of delay, the IOT compressed sensing and information interaction technology reduce 8.2% and 8.3%, respectively, compared with the traditional technology; in terms of throughput, the IOT compressed sensing and information interaction technology increase 8.1% and 12.2%, respectively, compared with the traditional technology.

The reason for the outstanding effect of Internet of things compressed sensing and information interaction technology is that the initial value of the competition window is set more reasonably after the message is sent successfully. When the number of vehicles is small, when the backoff algorithm is implemented, the initial value of backoff counter is small and the backoff time is short; when the number of vehicles is increased, the initial value of backoff counter is large and the backoff time is long. This method can adaptively adjust the initial value, change the value of the backoff counter according to the vehicle density, and reduce the possibility of channel transmission conflict as much as possible. Through the simulation, it is found that the different number of vehicle nodes also has an impact on the performance parameters of intelligent transportation layout. Therefore, this experiment designs an intelligent transportation layout environment with different number of vehicle nodes. The simulation results in Figure 7 show that, with the increase of the number of vehicle nodes, there will be a relatively good peak of its layout performance. This is because when the number of vehicles is small, the path to the destination node is relatively small, and the amount of information interaction data is relatively small, so the delay may be large, and the throughput is not high; with the increase of the number of vehicles, the routing establishment is more perfect, and the information interaction process makes full use of the intelligent transportation layout resources, so that the performance of message delay and throughput is the best. When the number of vehicles continues to increase, the resources of its layout will be relatively congested, resulting in the deterioration of its performance index again.

5.2. Validation Analysis of Simultaneous Interpreting Scenarios with Different Transmission Distances. In vehicle road cooperative system, the transmission distance of vehicle node is its communication radius. The transmission distance is directly related to the hop number of node information transmission and the number of neighbour nodes and also directly affects the strength of channel competition between different nodes. This simulation experiment compares the impact of optimized IOT compressed sensing and information interaction technology and traditional technology on the system communication performance under two different communication transmission distances of 100 m and 250 m. 250 m is the standard case. In this simulation experiment, the simulation results for message delay and throughput are 100 m and 250 m, respectively. When the transmission distance of vehicle nodes is 100 m and 250 m, respectively, in terms of delay, the Internet of things compressed sensing and information interaction technology protocols reduce

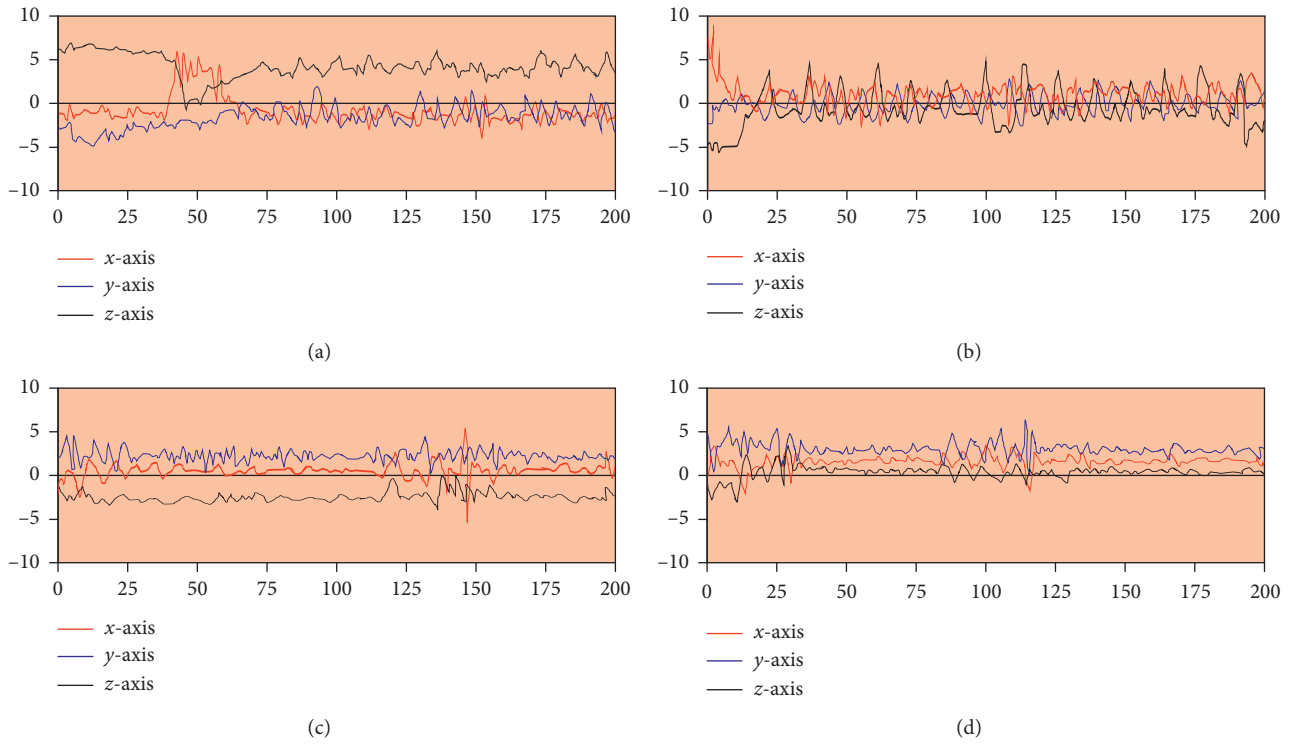


FIGURE 5: Decomposition of vehicle sensing speed.

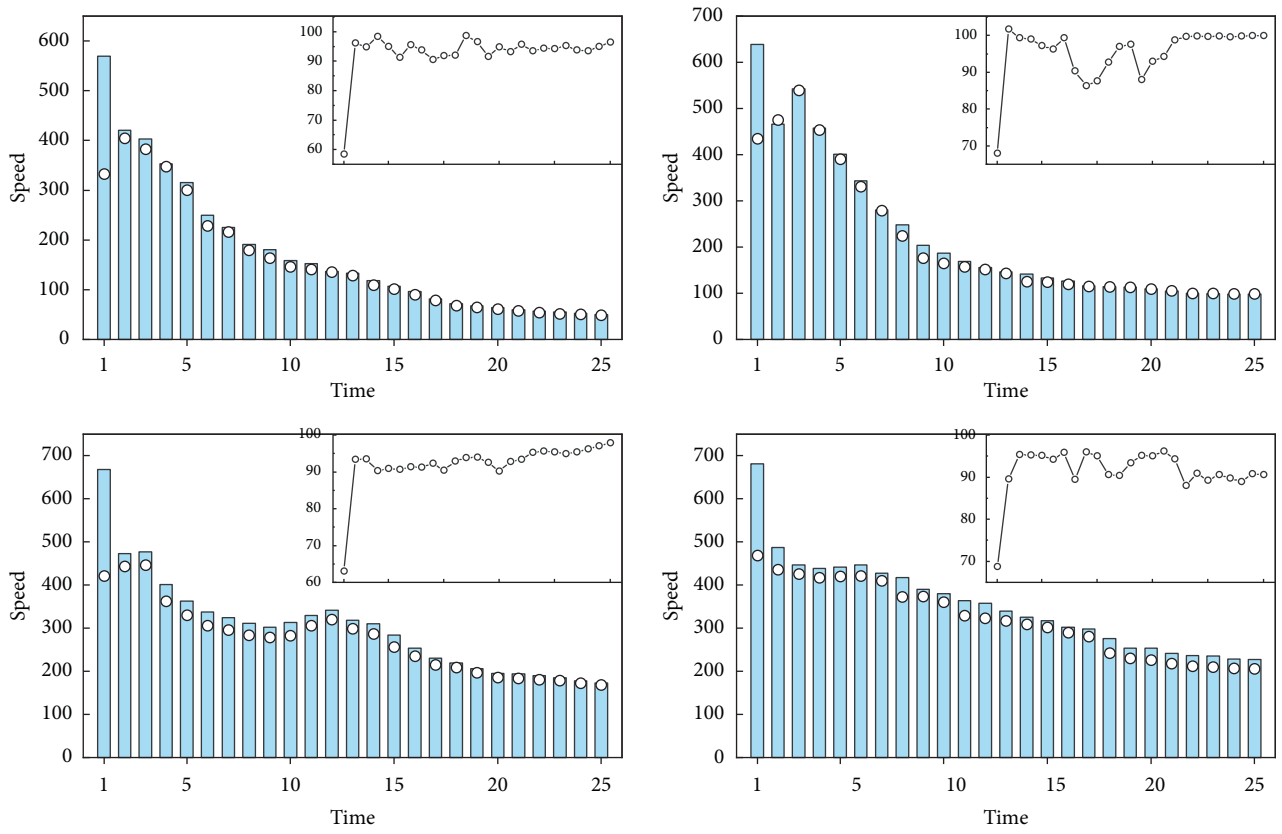


FIGURE 6: Compressed sensing reconstruction of vehicle sensing speed.

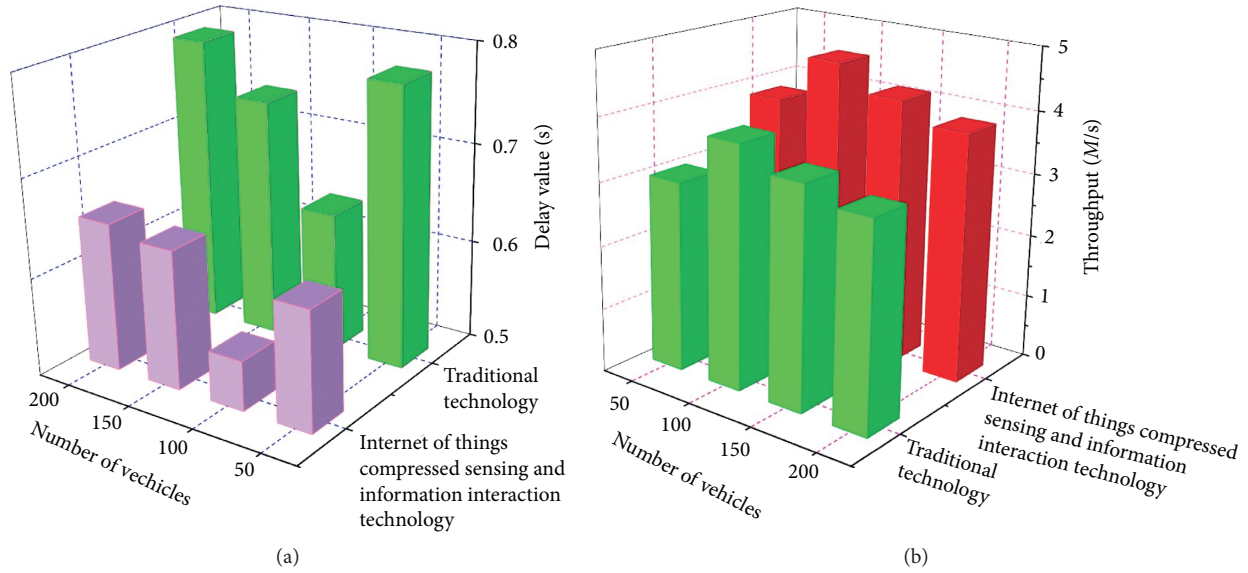


FIGURE 7: Simulation results of different vehicle density. (a) Comparison of time delay. (b) Comparison of throughput.

7.6% and 8.2%, respectively, compared with the traditional technology protocols, and in terms of throughput, the Internet of things compressed sensing and information interaction technology increase 14.1% and 8.1%, respectively, compared with the traditional technology protocols.

This experiment simulates the communication intelligent transportation layout in the case of multiple transmission distances, and the simulation results are shown in Figure 8. With the increase of transmission distance, the performance of its layout changes in a “U” shape. The reason for this phenomenon is that when the transmission distance is short and the number of transmission hops increases, the message delay is bound to be large and the throughput is not large. When the transmission distance is very large, the number of nodes communicating at the same time will decrease rapidly, the utilization rate of intelligent transportation layout resources will decrease, and the messages cannot be sent out in time, resulting in the increase of transmission delay and the decrease of intelligent transportation layout throughput. When the node transmission distance is relatively small, the probability of collision between nodes increases, resulting in an increase in the number of retransmissions; when the node transmission distance is relatively large, the number of transmission hops from the source node to the destination node increases, which takes up a long time of intelligent transportation layout resources, and will inevitably cause unnecessary transmission delay. Compared with the traditional technology, IOT compressed sensing and information interaction technology can sense the communication status and retransmission times of neighbour nodes and set reasonable backoff time slot.

5.3. Validation Analysis under Different Driving Speed Scenarios. The speed of vehicle nodes in vehicle road cooperative system has a decisive influence on the change rate of

intelligent transportation layout topology. When the vehicle speed is extremely fast, it may lead to the misjudgement of hidden nodes and exposed nodes. This simulation experiment designs and compares the impact of IOT compressed sensing and information interaction technology and traditional technology on the communication performance of vehicle road cooperative system under two different driving speeds of 20 km/h and 40 km/h and sets 40 km/h as the standard case. The vehicle speed plays a decisive role in the change of its topology and has a great impact on its layout performance. In general, the compressed sensing and information interaction technology protocol of the Internet of things has improved in terms of delay and throughput compared with the traditional technology protocol. In the case of 20 km/h and 40 km/h, the message delay is reduced by 8.5% and 8.2%, respectively; in terms of throughput, it is increased by 16.6% and 8.1%, respectively.

Through the simulation of the intelligent transportation layout of the vehicle road collaborative system under different vehicle speeds, it is found that the performance of the intelligent transportation layout is irregular, and there will not be peaks or troughs like the previous two experimental scenarios. The simulation results are shown in Figure 9. Speed is equivalent to random quantity in the vehicle road cooperative system. After all, compared with the radio transmission speed, the vehicle speed of 10 km/h–70 km/h in this simulation system can be ignored.

In general, the effect of IOT compressed sensing and information interaction technology protocol is still better than that of traditional technology, mainly because IOT compressed sensing and information interaction technology protocol adaptively adjusts the competitive time window according to the number of retransmissions of nodes, reduces the time of nodes waiting for the channel when sending conflicts, and orderly uses the wireless intelligence in the traffic environment which can be used to arrange transportation resources.

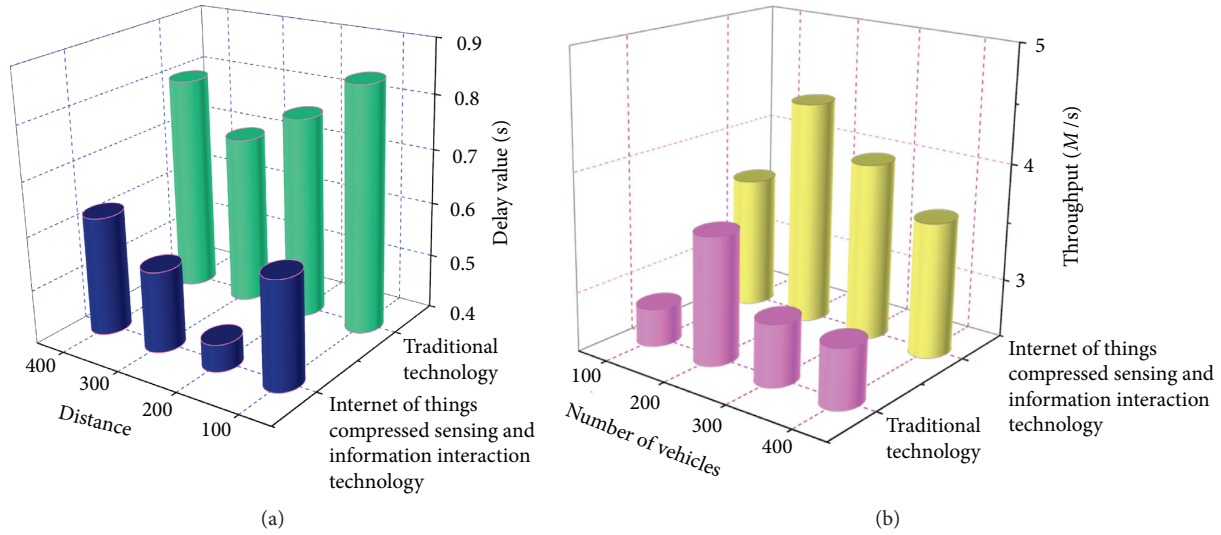


FIGURE 8: Simultaneous interpreting results of different transmission distances. (a) Comparison of time delay. (b) Comparison of throughput.

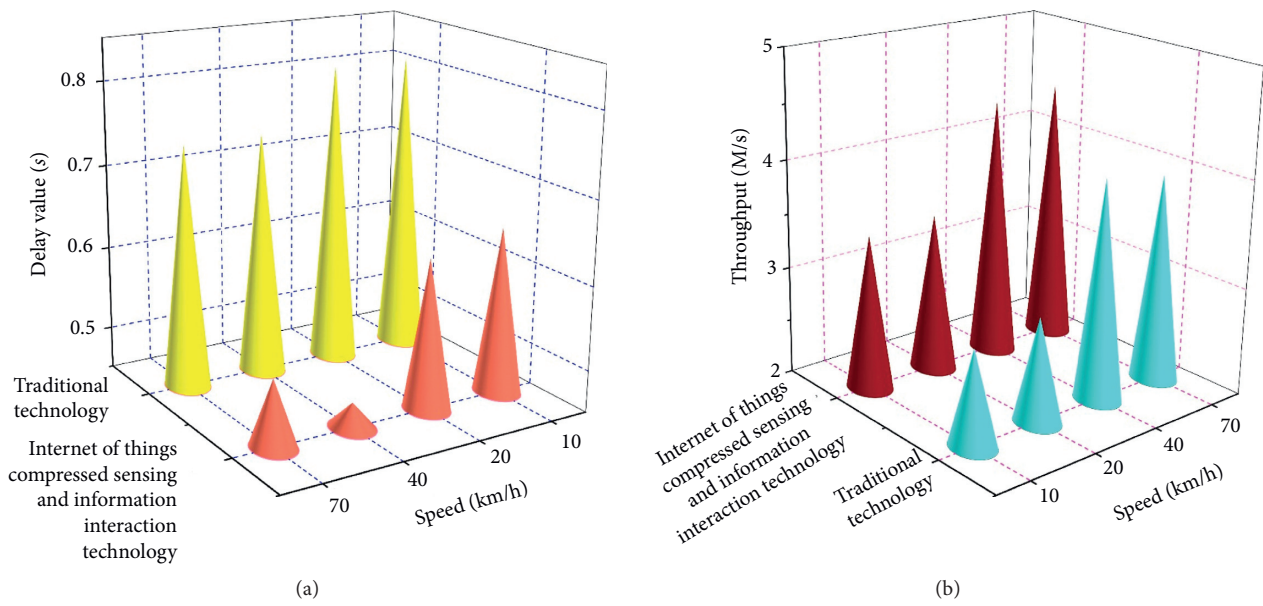


FIGURE 9: Simulation results of different vehicle speeds. (a) Comparison of time delay. (b) Comparison of throughput.

6. Conclusion

The rapid growth of car ownership has brought convenience to all aspects of people and led to frequent urban road congestion. Road congestion can have many negative effects. There are two main solutions, one is the traditional way of adding roads and controlling the increase in the number of vehicles, and the other is dredging. In the current situation, although the expansion of roads and the control of the increase in the number of vehicles can alleviate the traffic pressure, the utilization of land resources is limited after all. This way cannot solve the congestion phenomenon from the perspective of sustainable development. This paper describes the research status of intelligent transportation layout and

uses the basic theory of compressed sensing and information interaction technology to carry out data fusion and reconstruction algorithm in intelligent transportation. Finally, the application of compressed sensing in intelligent transportation layout is summarized to realize data collection in intelligent transportation network. In this paper, compressed sensing theory is applied to the data acquisition of intelligent transportation network to reduce the amount of data and improve the effective data acquisition. The reconstruction algorithm is used to realize the data reconstruction to ensure the accuracy and stability of signal reconstruction and the accuracy of network transmission information. Finally, the simulation analysis is given. But at present, the feature dimension is very large, and the

recognition process takes a long time, which cannot meet the real-time requirements. Further research is needed to guide the selection of feature dimension and the extraction of line features extracted by transformation and to solve the nonlinear pattern classification problem which is difficult to be solved by current compressed sensing methods. There will be more problems to be solved in the process of applying IOT compressed sensing and information interaction technology to the actual intelligent traffic layout.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Y. Gong and J. H. Liao, "Urban Intelligent Transportation big data platform based on blockchain technology and simulation case analysis," *Highway Transportation Technology*, vol. 36, no. 12, pp. 117–126, 2019.
- [2] D. Wu and J. W. Hu, "Information perception and interaction technology of Internet of things," *Computer Knowledge and Technology*, vol. 13, no. 27, pp. 225–226, 2017.
- [3] X. D. Wang, "Research on information perception and interaction technology of Internet of things," *Journal of Shanxi Coal Management Cadre College*, vol. 29, no. 1, pp. 204–205, 2016.
- [4] Y. Wu, "Research on information perception and interaction technology of Internet of things," *Journal of Chifeng University*, vol. 32, no. 20, pp. 36–37, 2016.
- [5] Y. D. Zhang and S. Z. Peng, "Construction of intelligent transportation system based on big data of internet of vehicles," *Comprehensive Transportation*, vol. 40, no. 11, pp. 25–29, 2018.
- [6] Y. Q. Xia, C. Yan, X. Q. Wang, and X. H. Song, "Intelligent transportation information physical fusion cloud control system," *Acta Automatica Sonica*, vol. 45, no. 1, pp. 132–142, 2019.
- [7] P. Zhang, "Application of internet of things technology in intelligent transportation," *Science and Technology Communication*, vol. 11, no. 4, pp. 158–159, 2019.
- [8] J. Sun, "Cross integration of artificial intelligence and intelligent transportation," *Information and Computer*, vol. 11, no. 4, pp. 139–141, 2019.
- [9] Y. S. Yang and X. Y. Li, "Research on intelligent transportation management and application based on big data technology," *Journal of Chongqing Business University*, vol. 36, no. 2, pp. 73–79, 2019.
- [10] Z. Y. Luo, "Application of artificial intelligence technology in urban intelligent transportation," *Communication World*, vol. 26, no. 5, pp. 244–245, 2019.
- [11] Z. Q. Yu and J. L. Li, "Design of intelligent transportation integrated information application platform based on big data," *Journal of Henan Normal University*, vol. 47, no. 4, pp. 37–41, 2019.
- [12] Y. Y. Zhang, "Analysis on the development and application of artificial intelligence in intelligent transportation," *Information Recording Materials*, vol. 20, no. 5, pp. 105–106, 2019.
- [13] W. Zhang, Y. Q. Wei, T. T. Liu, and S. L. Guan, "Application and platform construction of intelligent transportation system under the background of big data," *Shanxi Architecture*, vol. 45, no. 11, pp. 183–185, 2019.
- [14] L. J. Zeng, "Research on urban intelligent transportation system design based on big data," *Think Tank Era*, vol. 35, no. 7, pp. 281–282, 2019.
- [15] H. Mi and Y. Zheng, "Design and construction of intelligent transportation collaborative innovative training base based on intelligent transportation industry chain," *Journal of Jinling University of Science and Technology*, vol. 35, no. 3, pp. 26–31, 2019.
- [16] Y. Tao, X. D. Yan, T. M. Wang, and M. Liu, "Development strategy of intelligent transportation system for future intelligent society," *Science and Technology Guide*, vol. 34, no. 7, pp. 48–53, 2016.
- [17] L. B. Huang, "Review of urban intelligent transportation system construction at home and abroad," *Urban Roads, Bridges and Flood Control*, vol. 17, no. 5, pp. 40–45, 2016.
- [18] Y. Wu, "Application of internet of things technology in intelligent transportation system architecture," *Automation and Instrumentation*, vol. 20, no. 6, pp. 131–132, 2016.
- [19] F. Wang, D. S. Jiang, and H. M. Yue, "Research and development of intelligent traffic lights to effectively relieve urban traffic pressure," *Journal of Changchun Institute of Technology*, vol. 17, no. 3, pp. 101–104, 2016.
- [20] Z. H. Wu and Y. He, "Making full use of intelligent transportation technology to improve road traffic safety," *Traffic Information and Safety*, vol. 33, no. 1, pp. 1–8, 2015.
- [21] L. L. Bai and T. P. Han, "Application of big data in intelligent transportation system," *Computer Knowledge and Technology*, vol. 11, no. 10, pp. 204–206, 2015.
- [22] H. P. Lu, Z. Y. Sun, and W. C. Qu, "Big data and its application in urban intelligent transportation system," *Transportation System Engineering and Information*, vol. 15, no. 5, pp. 45–52, 2015.
- [23] H. F. Liu and P. H. Huang, "Intelligent transportation Internet of things technology and industrialization," *Engineering Research Engineering in Interdisciplinary Perspective*, vol. 6, no. 1, pp. 20–30, 2014.
- [24] M. J. Jin, "Development status and prospect of intelligent transportation system technology in China," *Traffic Information and Safety*, vol. 30, no. 5, pp. 1–5, 2012.
- [25] Z. S. Jiang, X. N. Tang, and X. H. Chen, "Architecture analysis of internet of vehicles and its application in intelligent transportation system," *Internet of Things Technology*, vol. 2, no. 11, pp. 39–41, 2012.
- [26] J. F. Bai, "Feasibility study on the development of urban intelligent transportation network system in the era of big data," *Communication World*, vol. 27, no. 1, pp. 157–158, 2020.
- [27] Y. Z. Liang, "System design of intelligent transportation data analysis platform under big data technology environment," *Value Engineering*, vol. 39, no. 15, pp. 243–244, 2020.

Research Article

Tumor Grade and Overall Survival Prediction of Gliomas Using Radiomics

Jianming Ye ¹, He Huang ², Weiwei Jiang ², Xiaomei Xu ², Chun Xie ¹, Bo Lu ³,
Xiangcai Wang ¹ and Xiaobo Lai ²

¹The First Affiliated Hospital, Gannan Medical University, Ganzhou 341000, China

²School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou 310053, China

³Faculty of Engineering, Shanghai Normal University Tianhua College, Shanghai 201815, China

Correspondence should be addressed to Bo Lu; lb2364@sthu.edu.cn, Xiangcai Wang; wangxiangcai@csc.ac.cn, and Xiaobo Lai; dmia_lab@zcmu.edu.cn

Received 19 March 2021; Revised 3 April 2021; Accepted 10 April 2021; Published 23 April 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Jianming Ye et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Glioma is one of the most common and deadly malignant brain tumors originating from glial cells. For personalized treatment, an accurate preoperative prognosis for glioma patients is highly desired. Recently, various machine learning-based approaches have been developed to predict the prognosis based on preoperative magnetic resonance imaging (MRI) radiomics, which extract quantitative features from radiographic images. However, major challenges remain for methodologic developments to optimize feature extraction and provide rapid information flow in clinical settings. This study investigates two machine learning-based prognosis prediction tasks using radiomic features extracted from preoperative multimodal MRI brain data: (i) prediction of tumor grade (higher-grade vs. lower-grade gliomas) from preoperative MRI scans and (ii) prediction of patient overall survival (OS) in higher-grade gliomas (<12 months vs. > 12 months) from preoperative MRI scans. Specifically, these two tasks utilize the conventional machine learning-based models built with various classifiers. Moreover, feature selection methods are applied to increase model performance and decrease computational costs. In the experiments, models are evaluated in terms of their predictive performance and stability using a bootstrap approach. Experimental results show that classifier choice and feature selection technique plays a significant role in model performance and stability for both tasks; a variability analysis indicates that classification method choice is the most dominant source of performance variation for both tasks.

1. Introduction

Glioma is one of the most common and deadly malignant brain tumors originating from glial cells. About 50 percent of nervous system tumors and 80 percent of all malignant brain tumors are gliomas. Glioblastoma multiforme (GBM) (also called glioblastoma) is a fast-growing glioma that develops from star-shaped glial cells (astrocytes and oligodendrocytes) that support the health of the nerve cells within the brain. In adults, GBM occurs most often in the cerebral hemispheres, especially in the brain's frontal and temporal lobes of the brain. GBM is a devastating brain cancer that typically results in death in the first 15 months after diagnosis. Traditional treatment of GBM is surgical resection followed by radiation therapy and/or chemotherapy.

However, the median survival time of GBM is still less than 15 months despite surgical resection, radiotherapy, and chemotherapy. Therefore, the accurate preoperative prognosis of GBM patients is desired, which can provide essential information for planning the optimized and personalized treatment.

Recently, various machine learning-based approaches have been developed to predict the prognosis based on preoperative magnetic resonance imaging (MRI) radiomics, which is a new cross-field of medical informatics, aiming to extract quantitative features defined by mathematics from medical images, such as shape, intensity, and texture [1, 2]. Particularly, they applied the regression model to predict OS time in days or categorized it into short or long term based on binary classification using radiomic

features extracted from various types of preoperative image data [3]. According to the strategy through which features are extracted, these studies can be roughly divided into two categories: (1) methods based on manual features and (2) methods based on automatically extracted features using machine learning techniques. The basic idea of the method based on manual features is to extract the artificial designed features by semiautomatic or full-automatic method and use the traditional machine learning method to regress or classify the calculated features [4, 5]. For example, in [6], brain tumors' image phenotypic features are calculated from a public preoperative multimodal MRI brain dataset and input into the random forest classifier to learn a regression model for OS prediction. In [7], some manually labeled features are extracted from the BraTS 2017 dataset, such as the volume and surface irregularity of brain tumors, are used to train the artificial neural networks for OS prediction. Although manual feature-based methods have shown promising results, there is no systematic way to determine OS-related manual features but mostly depended on experience. Therefore, the machine learning-based methods have been proposed, which can automatically learn OS-related, deeply embedded MRI image features to better predict OS without prior knowledge [4]. For example, in [8], Nie et al. present a two-stage deep learning-based OS time prediction method of high-grade gliomas patient, where a 3D multichannel convolutional neural network (CNN) is proposed to extract implicit and high-level features automatically for OS prediction from multimodal preoperative MRI brain tumor data, including the contrast-enhanced T1 (T1c), diffusion tensor imaging (DTI), and resting-state functional MRI (rs-fMRI). In [9], a novel three-dimensional detailed delineation algorithm is introduced for GBM tumors in MRI, which efficiently delineates the whole tumor, enhancing core, edema, and necrosis volumes using fuzzy connectivity and multi-thresholding, followed by survival prediction of patients using the concept of habitats.

Although all the studies mentioned above have indicated an essential value of brain imaging phenotype for OS prediction, tumors are often heterogeneous in space and time. There are differences in the cell, gene, and microenvironment for different tumor regions at the same time point or at other time points in the same tumor region, which usually requires multiple biopsies to capture the tumor's molecular heterogeneity, bringing inconvenience and risk to patients. Radiomics can provide a noninvasive way to explore the heterogeneity of tumors [10]. Gliomas are the most common primary malignant brain tumors with high intrinsic heterogeneity. This heterogeneity is evident in radiomic features and morphology, making classification and prognosis more difficult [11]. Radiomics analysis of gliomas can provide additional information about the patient's classification, prognosis, and possible survival outcomes [12, 13].

However, although researchers at home and abroad have done a lot of research on the application of machine learning algorithm in radiomic feature classification and prognosis prediction [14–19], due to the lack of a unified standard, there are still many unknowns about which is the

best model in the field. Many studies also use proprietary or in-house software in their radiomic feature extraction/analysis pipeline, severely limiting the community from making advances. Coupled with the fact that patients' medical images are protected by the confidentiality laws, it is incredibly challenging, if not impossible, to reproduce the results. Therefore, it is crucial to utilize publically available datasets and open-source tools to expand the radiomics field.

In our study, two machine learning classification tasks using radiomic features are investigated, which predict tumor grade and patient OS from preoperative MRI scans, respectively. These two tasks utilize the conventional machine learning techniques constructed with various classifier methods. Feature reduction methods also are applied to increase model performance and decrease computational costs. Models are assessed in terms of their predictive performance and stability using a bootstrap approach based on the 2017 BraTS Challenge's MRI data. Experimental results show that the classifier choice and dimensionality reduction technique plays a significant role in model performance and stability for both tasks. Figure 1 shows an outline of the radiomic workflow for the grade classification task, and we utilized a similar scheme for the overall survival classification task.

2. Material and Method

2.1. Dataset and Preprocessing. We utilized the 2017 BraTS Challenge's Training Dataset [20], which comprises 210 higher-grade gliomas (HGG) and 75 lower-grade gliomas (LGG) preoperative multimodal MRI scans collected from multiple centers. Each patient's multimodal scans include T1, postcontrast T1-weighted (T1c), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR). All the MRI scans have been segmented manually by one to four raters, following the same annotation protocol, and experienced neuroradiologists approved their annotations. Annotations comprise the GD-enhancing tumor (ET-label 4), the peritumoral edema (ED-label 2), and the necrotic and non-enhancing tumor (NCR/NET-label 1). Each sequence was skull-stripped and was resampled to $1\text{ mm} \times 1\text{ mm} \times 1\text{ mm}$ (isotropic resolution). For the overall survival challenge, age and prognosis of the patient posttreatment were supplied by the organizers. The overall survival data also were available for a subset of the GBM scans. In this study, all samples used for the grade prediction task are referred to as the Tumor Grade Dataset; all samples used for the overall survival prediction classification task are referred to as the Overall Survival Dataset. For the Tumor Grade Dataset, glioblastoma multiforme (GBM) was considered the negative type ($n=210$) while LGG was considered the positive type ($n=75$). The Overall Survival Dataset was stratified into binary classes based on median survival rates for GBM; patients who died before 12 months from diagnosis were considered negative ($n=81$), while patients who died after 12 months were considered positive ($n=82$). Examples of the four modalities and the corresponding tumor masks from two GBM patients are shown in Figure 2.

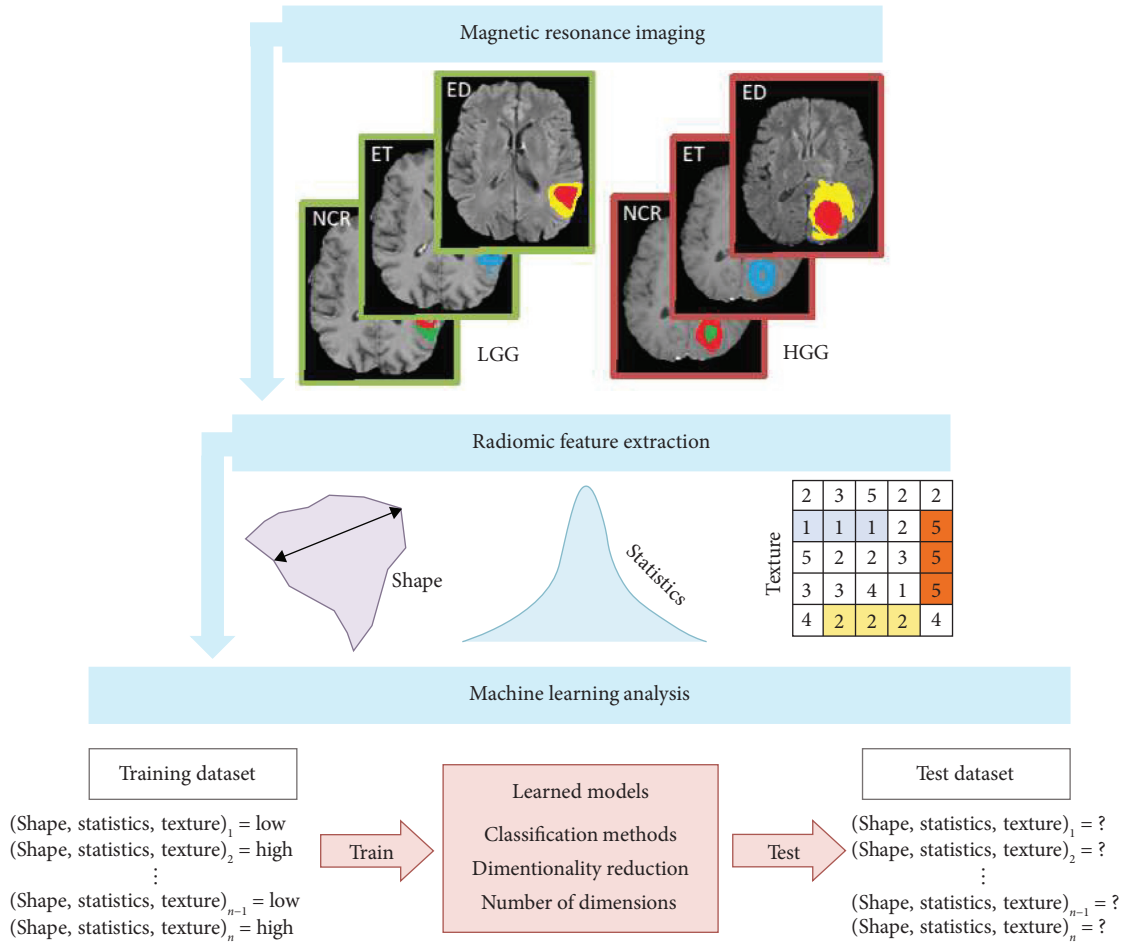


FIGURE 1: Proposed workflow for grade/survival classification task.

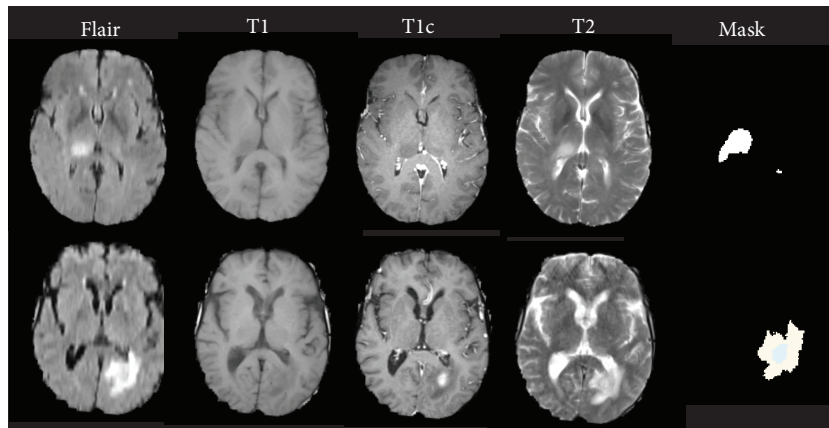


FIGURE 2: Examples of the four MRI modalities and the corresponding tumor masks from two randomly selected GBM patients.

2.2. Multitask VNet for Glioma MRI Data Segmentation.

This study uses a multitask VNet framework to segment glioma and its different subregions from the multimodal MR image, shown in Figure 3. The network has two decoder modules with similar structures, and different decoders are assigned different tasks. The mask decoder module performs training-mask segmentation according to pixel classification tasks, and the distance transform decoder module performs

regression tasks to realize distance map estimation. The structure of the encoder module and decoder module of the network is similar to the VNet. Its encoder module alternately stacks convolutional layers and downsampling layers to achieve feature extraction of the input signal under different receptive fields.

In contrast, the decoder module alternately stacks deconvolutional layers and convolutional layers in the joint

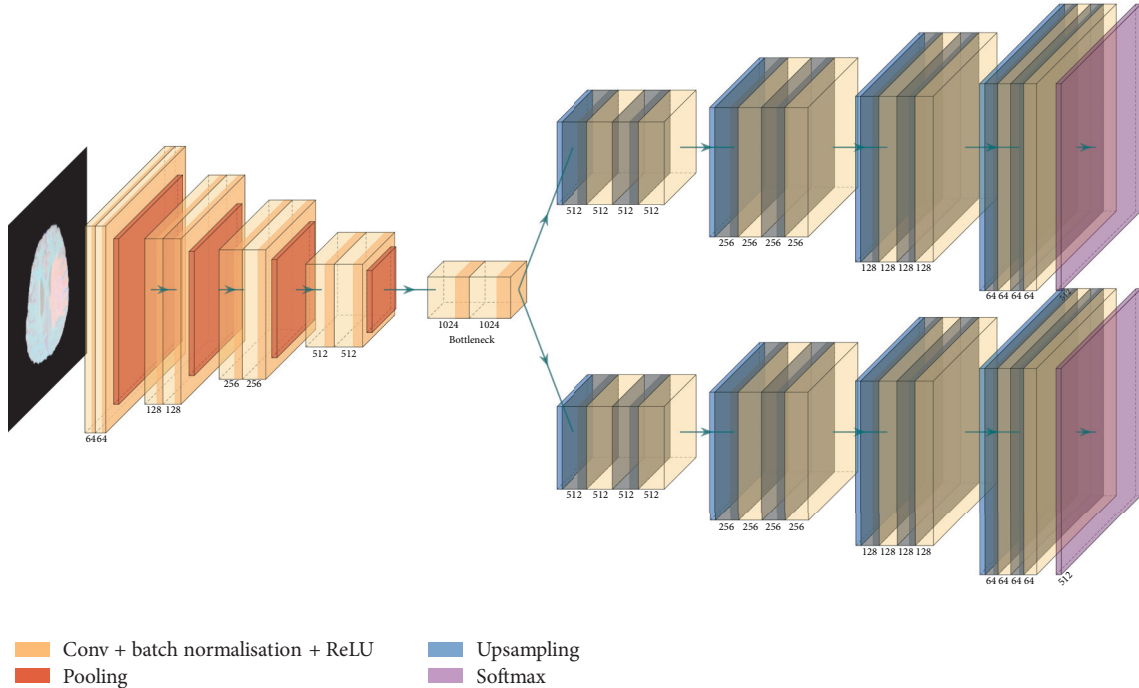


FIGURE 3: The overall architecture of the proposed multitask VNet.

encoder restore image resolution stage by stage based on the features extracted by the module. The model's loss function is the weighted sum of the categorical focal loss of the mask decoder block and the MSE loss of the distance transform decoder block. Its essence is that the distance map prediction regularizes the template prediction.

2.3. Quantitative Feature Extraction. Medical images contain a lot of information that can reflect the relationship between human macro performance and microenvironment. Up to now, the analysis and diagnosis of medical images are mainly based on human judgment. The disadvantage of this method is that it can only be qualitative but not quantitative. Compared with the qualitative description of human experience, quantitative features can reflect more potential information in the image. Medical imaging has developed from traditional morphological diagnosis to quantitative tumor analysis. The main difference is that the latter needs to extract and analyze more high-order quantitative image features.

Quantitative feature extraction refers to the process of extracting information from images by computer. The performance of a classification model largely depends on the features used. We extracted 16 shapes, 19 first-order statistics, 27 gray-level cooccurrence matrix (GLCM), 16 gray-level size zone matrix (GLSZM), and 16 gray-level run length matrix (GLRLM) features from each phenotype region of interest (ROI). The coiflet wavelet transform filter was also applied to each image to extract eight decompositions; for each phenotype, each decomposition's intensity-based features were calculated. The combination of shape features, first-order features, texture features, and wavelet features extracts 718 features for each image phenotype and 2154

features for each sample. Before extracting these features, voxel intensity values were normalized using the Z-score normalization in the whole brain, discretized with a bin width of 0.1, and constrained to an intensity value range of 3 standard deviations from the mean. For the Tumor Grade Dataset, some LGG samples do not contain ET segmentations. Therefore, regardless of tumor grade, these samples were removed from the analysis to keep the features equal. In addition, several mask combinations suffered from geometry mismatches and were likewise discarded. The removal of these samples from the Tumor Grade Dataset led to 44 LGG samples and 191 GBM samples remaining for the analysis. Similarly, after the removal of inappropriate samples, the Overall Survival Dataset was left with a total of 73 GBM samples with survival <12 months and 77 GBM samples with survival >12 months.

2.4. Feature Selection Methods. Radiomics leads to the creation of several informative features for use in predictive modeling. However, when the number of samples is far less than the number of features, direct classification prediction has a high computational cost and a poor effect. It may even lead to the classification prediction algorithm's failure. Hence, feature selection is needed to obtain the feature set with good performance after image feature extraction.

For machine learning models, there are many methods to reduce the feature space. Common categories of feature selection methods include filter, wrapper, and embedded methods. In addition, compared with the wrapper and embedded methods, the filter methods have the advantages of classifier independence and high computational efficiency [21]. Surprisingly, previous studies have found univariate filter methods that ignore interactions between variables can

be just as effective as multivariate methods that consider these interactions [22]. A possible alternative way for feature selection is dimensionality reduction. The complex interaction between variables is often considered by linear or nonlinear mapping, and the high-dimensional space is transformed into space with lower dimension [23]. It has been recently proposed that unsupervised dimensionality reduction techniques could have better prediction performance than filter methods in radiomic studies [24].

We utilized four unsupervised dimensionality reduction methods to build machine learning models, that is, principal component analysis (PCA), kernel PCA (KPCA), independent component analysis (ICA), and factor analysis (FA). We chose these methods due to their simplicity, computational efficiency, and easily available implementation. Moreover, these methods were compared with a univariate filter technique, ANOVA F -score with the top 30 features selected (FILT), and maximum 2D diameter features from each phenotype (DIAM). DIAM was chosen to investigate how our radiomic methods would compare against a commonly utilized prognostic radiological metric [25].

2.5. Models' Building. The prediction of tumor grade or overall survival in this paper is a small sample binary classification problem. To solve this problem, supervised learning in machine learning is more targeted. Supervised learning uses the training data to find rules through training to predict new samples. Training data consists of examples represented by a set of input features (radiomic features) and an output value (tumor grade or overall survival class). Once an intelligent prediction model is built from labeled data using a classifier and feature selection method, it can predict an unlabeled sample class.

We selected nine conventional machine learning techniques constructed with various classifier methods and two deep learning-based models for comparison, that is, decision trees (DT), random forest (RF), bagging (BAG), boosting (BST), Gaussian naïve Bayes (NB), multilayer perceptron (MLP), support vector machines (SVM), logistic regression (LR), k -nearest neighbors (KNN), convolutional neural networks (CNN), and deep neural networks (DNN). These models were chosen for their widespread use in radiomic studies and simple implementation. Models built with conventional machine learning and deep learning-based techniques are displayed in Table 1. Although the hyperparameter of classifiers can be tuned by cross-validation to improve the model performance, in our study, the classifiers were used together with the default hyperparameter settings to maintain simplicity and reduce the computational cost. Intelligent prediction models were built from combinations of feature reduction methods and classifier methods.

3. Experiment

3.1. Experimental Details. To analyze our results, a split was made by the patient. For each dataset (Tumor Grade Dataset $n=245$ and Overall Survival Dataset $n=150$), data were randomly split into training and testing sets with a test

size = 0.2, yielding training sets containing 196/120 samples and testing sets containing 49/30 samples, respectively. To prevent the class imbalances from affecting the models' performance, we applied the synthetic minority over-sampling (SMOTE) [26] technique to the tumor grade training dataset due to the existing radiomics studies that have shown SMOTE can effectively improve the classification predictive performance when the classes are imbalanced. However, SMOTE was not applied to the Overall Survival Dataset since classes were already balanced. Moreover, multicenter data and magnetic field inhomogeneities often contribute to the intensity inhomogeneities in the MR images. Therefore, we use the Z -score normalization as a necessary preprocessing step in dimensionality reduction to standardize features concerning the training set.

To investigate and compare the performance of different dimensionality reduction and classification approaches, a three-dimensional parameter grid for analysis was constructed in this study. For any of the four dimensionality reduction approaches, we took two as the step size ($n=1, 3, 5, 15$) incrementally selected the number of dimensions from 1 to 15 (e.g., principal component). The training data and 11 machine learning models evaluate these dimension subsets to build the machine learning prediction model. The area under the receiver operating curve (AUC) score was calculated to evaluate the model quantitatively on the test set, which was repeated 100 times for each combination with different random splits through a bootstrap approach. The mean of the AUC values (μ_{AUC}) over all iterations was calculated to determine the given model's final AUC value. By calculating the mean over 100 iterations, we can ensure a more representative value for each model. Similarly, an empirical metric for stability, relative standard deviation (RSD) was previously defined as follows [22]:

$$RSD = \frac{\sigma_{AUC}}{\mu_{AUC}}, \quad (1)$$

where σ_{AUC} and μ_{AUC} were the standard deviation and mean of the 100 AUC values, respectively. It should be noted that higher stability corresponds to lower RSD values.

We apply the popular open-source machine learning python library scikit-learn for model building and analysis in Python 3.6. The training and testing experiments are performed on an NVIDIA GeForce Titan RTX 24G GPU with Intel Xeon Silver 4210 2.2G GPU. The presented figures are generated using the plotting library Matplotlib. An open-source radiomics toolbox, Pyradiomics, was used for radiomic feature extraction.

3.2. Performance Measurements. There are three main experimental factors in our study which can affect the radiomics-based prediction, that is, prediction model (RF, NB, DT, BAG, BST, SVM, LR, MLP, KNN, CNN, and DNN), feature selection method (PCA, KPCA, ICA, and FA), and the number of dimensions selected (1, 3, 5, . . . , 15). Multivariate analysis of variance (ANOVA) was used to quantify these factors' impacts on AUC scores and their interactions in each classification task. To compare the variability contributed by

TABLE 1: Models built with various machine learning techniques.

Classifier methods	Dimensionality reduction methods	Feature selection methods
Decision trees (DT)	Principal component analysis (PCA)	ANOVA F -score (FILT)
Random forest (RF)	Kernel PCA (KPCA)	Max 2D diameter (DIAM)
Bagging (BAG)	Independent component analysis (ICA)	—
Boosting (BST)	Factor analysis (FA)	—
Naïve bayes (NB)	—	—
Multilayer perceptron (MLP)	—	—
Support vector machine (SVM)	—	—
Logistic regression (LR)	—	—
k -Nearest neighbor (KNN)	—	—

each factor, the variance (sum of squares) calculated for each factor was divided by total variance and multiplied by 100 to yield the percent variance for each factor.

In our study, a total of 2154 features were extracted from the segmented tumor regions of the preoperative MRI scans from the BraTS 2017 glioma dataset. For the Tumor Grade Dataset, the output classes were LGG or HGG, while for the Overall Survival Dataset, the output classes were <12-month or >12-month survival. For both classification tasks, feature selection and classification training were made using the training set, whereas the testing set was used to assess performance and stability.

3.2.1. Predictive Performance. Figure 4 depicts the performance of dimensionality reduction and classification methods using 11 dimensions for both tasks. Performance from models constructed using ANOVA F -score univariate filter method (FILT) and diameter features (DIAM) are also displayed. For the grade classification task, the best results among the four dimensionality reduction techniques are achieved by FA, while ICA usually performs the worst effects. Moreover, FA has comparable results to FILT, which generally has the highest predictive performance. Additionally, using diameter features alone scores much lower than any dimensionality reduction techniques. In terms of classifiers, most classifier methods show similar results except DT, which is noticeably lower. For the survival classification task, the best results among the four dimensionality reduction techniques are also often achieved by FA. Otherwise, performance results are more similar than in the grade classification task. Worthy of note is that using diameter features alone often scores comparable or higher than any dimensionality reduction techniques. Again, most classifier methods show similar results except decision trees and support vector machines (for PCA, KPCA, and ICA), which are noticeably lower. Additionally, AUC scores for the survival classification task are much lower (<0.65) than for the grade classification task (>0.80).

In addition, we repeated the above experiment by varying the number of dimensions. Figures 5 and 6 show the predictive performance corresponding to 1, 3, 5, 7, 9, 11, 13, and 15 dimensions for each feature selection method for both tasks, respectively.

3.2.2. Stability and Predictive Performance. Four AUC/RSD values corresponding to different dimensionality reduction techniques (PCA, KPCA, ICA, and FA) are generated for

each prediction method. We took the median of all four AUC/RSD values for each prediction task as the representative AUC/RSD of a model. Figure 7 shows the evaluation of models' representative stability and predictive performance in each classification task. In addition, as deep learning models show significantly performance in these tasks. Hence, the performance of both the conventional machine learning techniques and two deep learning-based models (DNN and CNN) is also evaluated. Figure 7(a) shows MLP, LR, KNN, and BST should be preferred as their stability and predictive performance were higher than the corresponding median values across all classifiers. Similarly, in Figure 7(b), MLP, LR, and KNN should be preferred, with BST on the borderline of top performance and stability.

3.2.3. Experimental Factor Effect. To quantify the effect of classification methods, dimensionality reduction methods, and the number of selected dimensions, multivariate ANOVA was performed on AUC scores in this study. In Figure 8, we observed that all three experimental factors and their interactions affect both classification tasks' prediction performance. The classification method was the most dominant source of variability as it explained 36% and 37% of the total variance in AUC scores for tumor grade and survival classification tasks, respectively. The number of dimensions used was the second most dominant source of variability for both tasks as it explained 28% and 20% of the total variance in AUC scores for tumor grade and survival classification tasks, respectively. The dimensionality reduction method was the least dominant source of variability for both tasks as it explained 3% and 2% of the total variance in AUC scores for tumor grade and survival classification tasks, respectively. Interaction terms between the experimental factors followed similar trends.

3.3. Discussion. Several studies have built radiomics-based predictive models for various clinical factors such as tumor grade, prognostic outcome, treatment response, and more. However, to expand the radiomics community, studies utilizing open-source data, tools, and machine learning models, such as those used in our current investigation, are necessary. In a series of papers by Parmar et al., they evaluated the predictive performance and stability of computed tomography (CT) radiomic machine learning models constructed with various feature selection filter

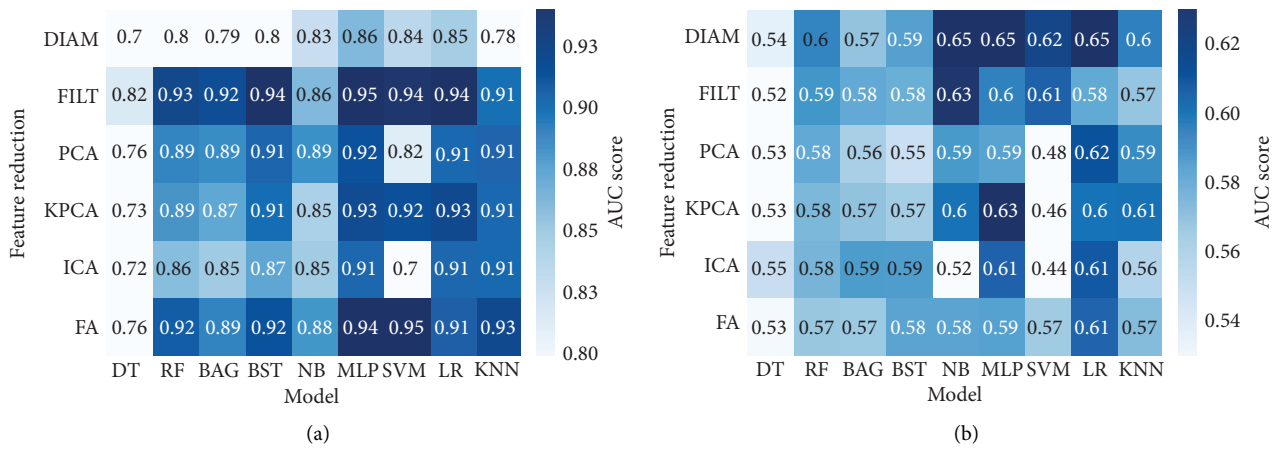


FIGURE 4: Predictive performance of feature reduction and classification methods. (a) Grade classification task and (b) survival classification task.

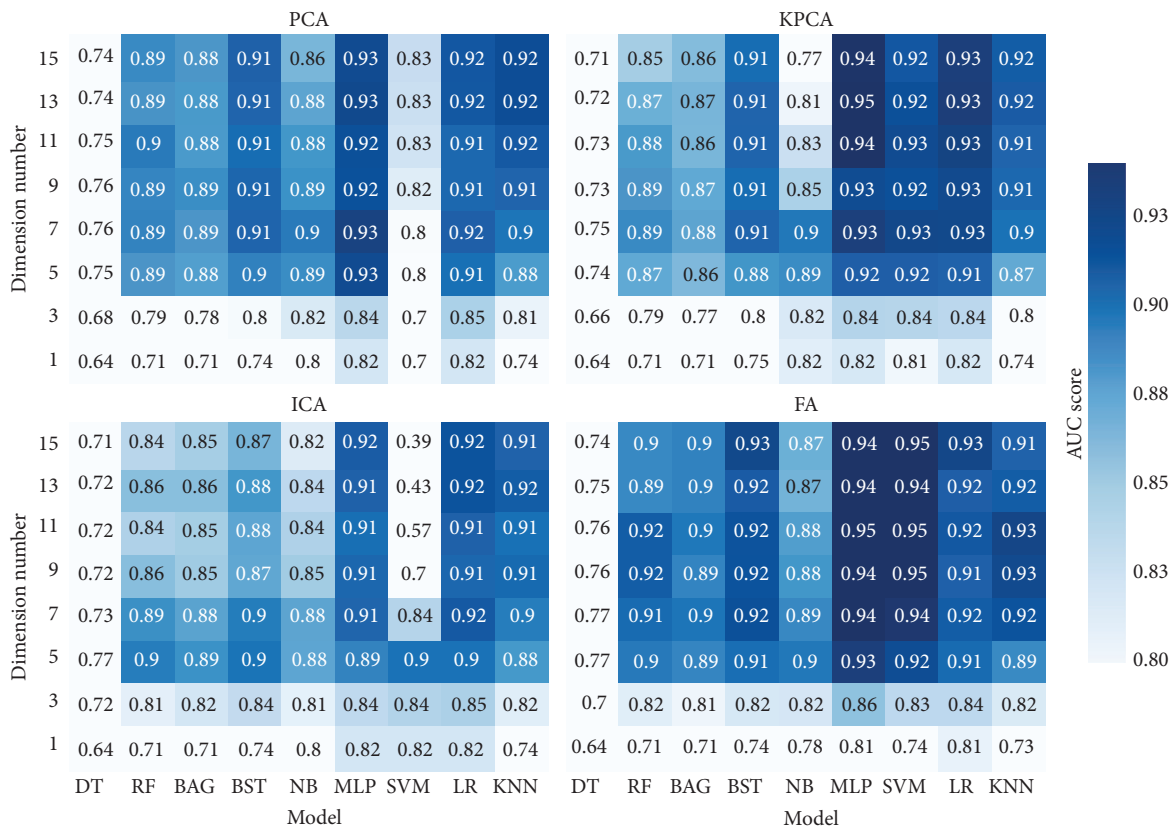


FIGURE 5: Predictive performance corresponding to classification methods and the number of dimensions for each dimensionality reduction method for grade classification task.

methods and classifier methods [22]. Results show that specific machine learning models perform differently depending on the cancer type, e.g., head and neck vs. lung. Therefore, it is vital to test these methods in different cancer types and various imaging modalities.

Additionally, Zhang et al. performed a similar study on lung CT with unsupervised dimensionality reduction methods and proposed dimensionality reduction methods have the potential to be superior to filter methods [27]. This

study further demonstrates the variability of machine learning models constructed from different classifiers and dimensionality reduction techniques in a different cancer type (glioma) and imaging modality (MRI). We demonstrate that dimensionality reduction techniques are often lower than or comparable to filtering methods for both tasks. Specifically, we show that FA can be an improvement over PCA, which was suggested by Zhang et al. to be the best method for dimensionality reduction in radiomic studies.

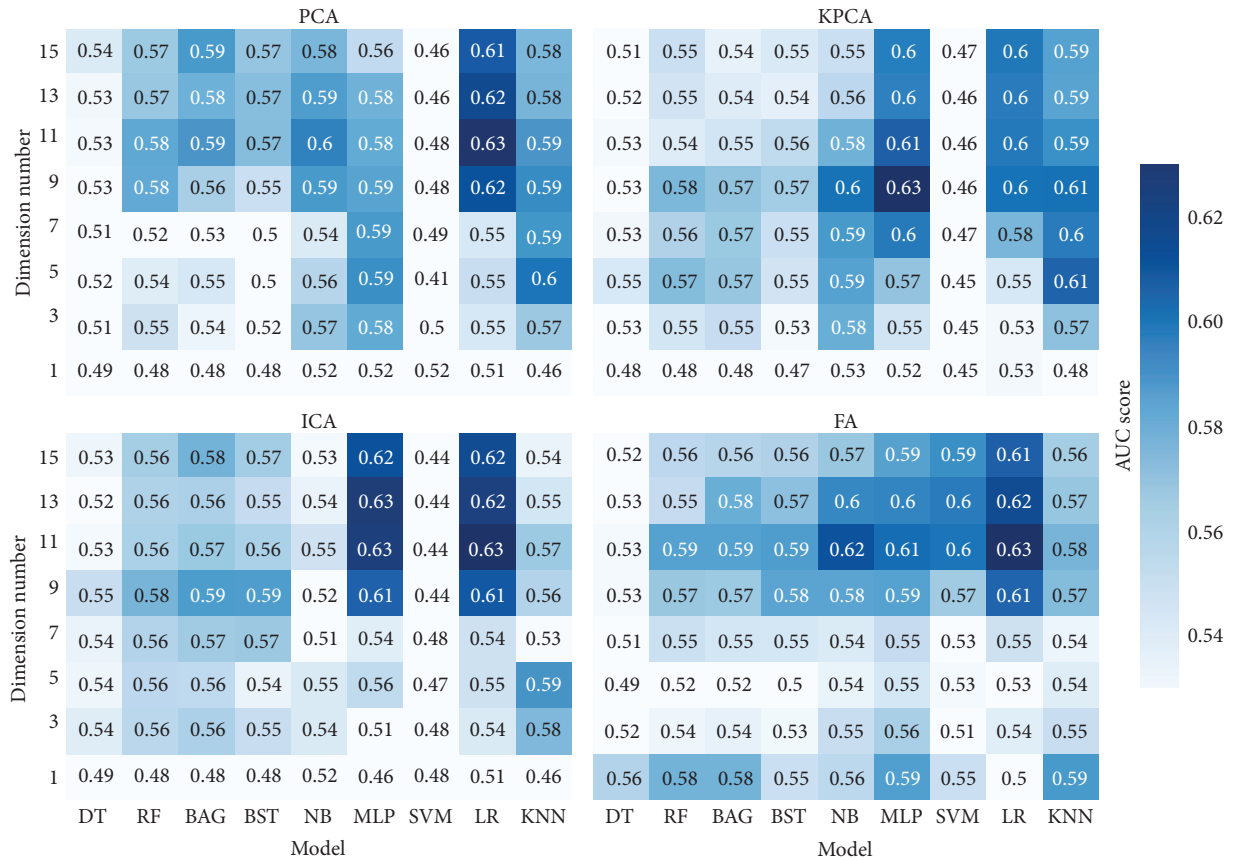


FIGURE 6: Predictive performance corresponding to classification methods and the number of dimensions for each dimensionality reduction method for survival classification task.

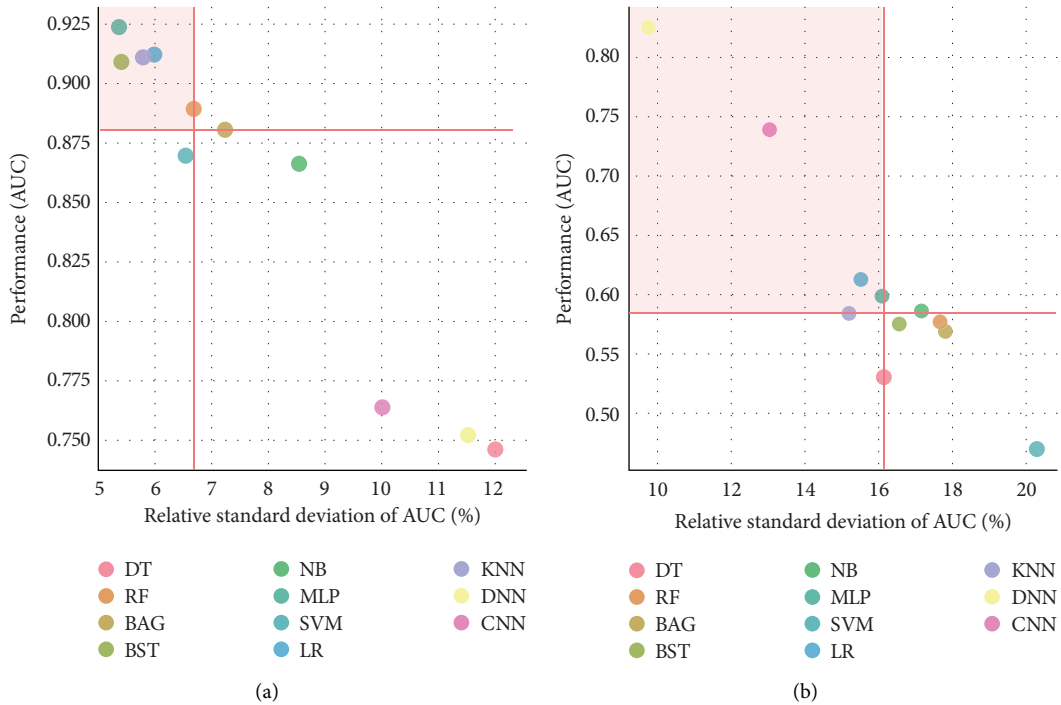


FIGURE 7: Scatterplots between representative stability and predictive performance of classification methods. (a) Grade classification task and (b) survival classification task.

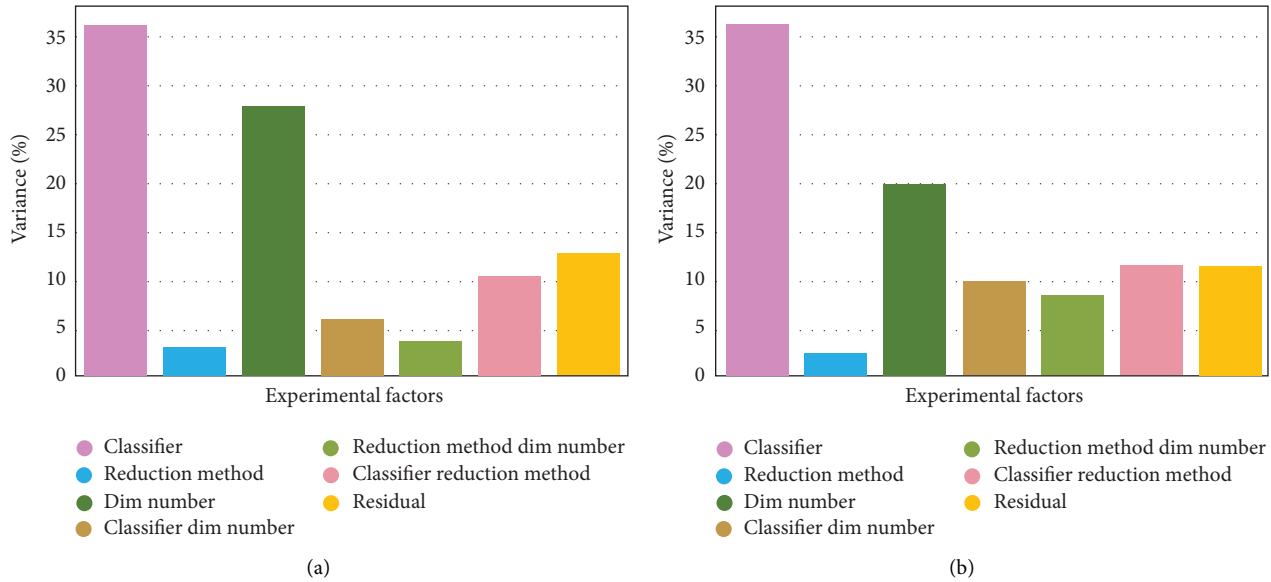


FIGURE 8: Variation of AUC explained by experimental factors and their interactions. (a) Grade classification task and (b) survival classification task.

ET in T1c MRI scans often used as a distinctive marker when attempting to distinguish LGG from HGG. However, since we have only used LGG samples that contain ET components, we suggest radiomics provides novel information about underlying phenotype, usually not possible in the radiological setting. Glioma grade is histopathologically diagnosed; i.e., a biopsy must be taken for classification [28]. With our radiomics approach, we suggest that imaging data may be a useful supplement to histological data. In this study, we have only classified LGG from HGG, but more grade subclasses can be assessed using these radiomics methods, e.g., grade 1 vs. grade 2 vs. grade 3 vs. grade 4. Previous studies have attempted to build machine learning models for glioma grade classification with dimensionality reduction techniques [29] or other feature selection methods [30]. Still, our results show higher predictive performance, possibly due to a more extensive training set and class balancing with SMOTE.

Predictive performance for grade classification is much higher when compared to survival classification, which is not surprising as each classification task has its own set of optimal radiomic biomarkers linked to underlying biological significance. For example, the combination of shape, first-order statistics, texture, and wavelet features utilized through dimensionality reduction leads to higher predictive performance than diameter features alone for the grade classification task. However, this is not the case for the survival classification task. Moreover, using diameter features alone in survival prediction leads to higher predictive performance than dimensionality reduction or filter techniques with all radiomic features. Previous studies have shown that texture features are challenging to gain predictive power from in GBM, with AUC values routinely falling <0.6 [17, 31]. It may be the case that current intensity-based features are not strongly linked to survival outcome in GBM, but further studies are necessary before coming to these

conclusions. This study has taken a coarse approach to build machine learning models, so it may very well be the case that more refined models for survival prediction can create useful texture-based radiomics signatures for GBM survival prediction with high AUC values.

For both classification tasks, the classifier method was the most significant contribution to variability in predictive performance. A trend has commonly been observed in radiomic studies investigating machine learning models using different classifiers and feature selection methods [22]. Oppositely, Wang et al. observed that the dimensionality reduction method plays a larger role in predictive performance variability [24]. Our study has also investigated the role the number of dimensions has in variability, and it was found that it has a larger role than the dimensionality reduction method used. To our knowledge, no other studies have investigated this factor's effect on predictive performance.

Some limitations of our study are as follows. Regarding image preprocessing, we have only utilized a simple method of intensity normalization (Z-score) due to its availability in Pyradiomics. Unlike CT imaging, MRI intensity is expressed in arbitrary units, necessitating intensity standardization before radiomic analysis. More sophisticated intensity normalization methods, such as histogram-based method [32], should be explored in future studies. In addition, we have not taken advantage of classifier hyperparameter tuning and instead relied on default hyperparameter settings to save on computational costs. Future studies should employ hyperparameter tuning to increase predictive performance and stability. While our research explores our classifiers' stability, it should be noted that RSD is only an empirical method that should not be directly compared with other studies but only as a relative reference between classifiers in a given study. Additionally, our definition of a top classifier is relative to other classifiers studied, so it should not be taken as all-encompassing.

4. Conclusion

In this study, we investigate two machine learning classification tasks using radiomic features: (i) prediction of tumor grade (higher-grade vs. lower-grade gliomas) and (ii) prediction of overall survival in higher-grade gliomas (<12 months vs. >12 months). These tasks are attempted using machine learning models constructed with various classifier methods and dimensionality reduction techniques. Models are assessed in terms of their predictive performance and stability using a bootstrap approach. Our results demonstrate that for both classification tasks, among dimensionality reduction methods, FA yielded the highest predictive performance. Similarly, MLP, LR, and KNN produced the highest predictive performance and stability among classifier methods. In addition, DT tended to perform poorly for both classification tasks. This possibly points to an underlying radiomic structure in the BraTS dataset that is preferentially fit by specific machine learning models. Where results start to diverge significantly is in the implementation of the SVM classifier. For the grade classification task, SVMs tend to perform relatively well with all feature selection methods except ICA. For the survival classification task, SVMs tend to perform poorly with all feature selection methods except FA. Interestingly, previous studies in different cancer types have suggested RF to be the best classifier method for radiomics studies. Still, it does not score among the best classifier methods for either task in our research.

Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was funded in part by the National Natural Science Foundation of China (Grant nos. 62072413 and 61602419), in part by the Natural Science Foundation of Zhejiang Province of China (Grant no. LY16F010008), in part by Medical and Health Science and Technology Plan of Zhejiang Province of China (Grant no. 2019RC224), and also in part by the Teacher Professional Development Project of Domestic Visiting Scholar in Colleges and Universities of Zhejiang Province of China (Grant nos. 2020-19 and 2020-20).

References

- [1] L. Xu, P. Yang, W. Liang et al., "A radiomics approach based on support vector machine using MR images for preoperative lymph node status evaluation in intrahepatic cholangiocarcinoma," *Theranostics*, vol. 9, no. 18, pp. 5374–5385, 2019.
- [2] Y. Wu, L. Xu, P. Yang et al., "Survival prediction in high-grade osteosarcoma using radiomics of diagnostic computed tomography," *EBioMedicine*, vol. 34, pp. 27–34, 2018.
- [3] Z. Tang, Y. Xu, L. Jin et al., "Deep learning of imaging phenotype and genotype for predicting overall survival time of glioblastoma patients," *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 2100–2109, 2020.
- [4] L. Liu, H. Zhang, J. Wu et al., "Overall survival time prediction for high-grade glioma patients based on large-scale brain functional networks," *Brain Imaging and Behavior*, vol. 13, no. 5, pp. 1333–1351, 2019.
- [5] B. Jie, D. Q. Zhang, W. Gao et al., "Integration of network topological and connectivity properties for neuroimaging classification," *IEEE Transactions on Bio-Medical Engineering*, vol. 6, no. 2, pp. 576–589, 2014.
- [6] B. H. Menze, J. Andras, B. Stefan et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Transaction on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.
- [7] A. Jungo, R. McKinley, R. Meier et al., "Towards uncertainty-assisted brain tumor segmentation and survival prediction," *International MICCAI Brainlesion Workshop*, vol. 25, pp. 474–485, 2017.
- [8] D. Nie, J. F. Lu, H. Zhang et al., "Multi-channel 3D deep feature learning for survival time prediction of brain tumor patients using multi-modal neuroimages," *Scientific Reports*, vol. 9, pp. 1–14, 2019.
- [9] S. Bhadani, S. Mitra, and S. Banerjee, "Fuzzy volumetric delineation of brain tumor and survival prediction," *Soft Computing*, vol. 24, no. 17, pp. 13115–13134, 2020.
- [10] H. J. W. L. Aerts, R. V. Emmanuel, T. H. Ralph et al., "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature Communications*, vol. 5, pp. 1–8, 2014.
- [11] D. J. Brat, R. G. W. Verhaak, K. D. Al-Dape et al., "Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas," *New England Journal of Medicine*, vol. 372, no. 26, pp. 2481–2498, 2015.
- [12] A. Kotrotsou, P. O. Zinn, and R. R. Colen, "Radiomics in brain tumors," *Magnetic Resonance Imaging Clinics of North America*, vol. 24, no. 4, pp. 719–729, 2016.
- [13] S. Narang, M. Lehrer, D. Yang, J. Lee, and A. Rao, "Radiomics in glioblastoma: current status, challenges and potential opportunities," *Translational Cancer Research*, vol. 5, no. 4, pp. 383–397, 2016.
- [14] J. Wang, C.-J. Wu, M.-L. Bao, J. Zhang, X.-N. Wang, and Y.-D. Zhang, "Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer," *European Radiology*, vol. 27, no. 10, pp. 4082–4090, 2017.
- [15] L. Zhang, Z. Ye, L. Ruan, and M. Jiang, "Pretreatment MRI-derived radiomics may evaluate the response of different induction chemotherapy regimens in locally advanced nasopharyngeal carcinoma," *Academic Radiology*, vol. 27, no. 12, pp. 1655–1664, 2020.
- [16] M. Ingrisich, M. J. Schneider, D. Nörenberg et al., "Radiomic analysis reveals prognostic information in T1-weighted baseline magnetic resonance imaging in patients with glioblastoma," *Investigative Radiology*, vol. 52, no. 6, pp. 360–366, 2017.
- [17] P. Grossmann, V. Narayan, K. Chang et al., "Quantitative imaging biomarkers for risk stratification of patients with recurrent glioblastoma treated with bevacizumab," *Neuro-Oncology*, vol. 19, no. 12, pp. 1688–1697, 2017.

- [18] B. Zhang, X. He, F. Ouyang et al., "Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma," *Cancer Letters*, vol. 403, no. 10, pp. 21–27, 2017.
- [19] P. P. J. H. Langenhuizen, S. Zinger, S. Leenstra et al., "Radiomics-based prediction of long-term treatment response of vestibular schwannomas following stereotactic radiosurgery," *Otology & Neurotology*, vol. 41, no. 10, pp. E1321–E1327, 2020.
- [20] J. Long, G. Ma, H. Liu et al., "Cascaded hybrid residual U-Net for glioma segmentation," *Multimedia Tools and Applications*, vol. 79, no. 33–34, pp. 24929–24947, 2020.
- [21] M. M. Rahman, O. L. Usman, R. C. Muniyandi et al., "A Review of machine learning methods of feature selection and classification for autism spectrum disorder," *Brain Sciences*, vol. 10, no. 12, pp. 1–23, 2020.
- [22] C. Parmar, P. Grossmann, J. Bussink et al., "Machine learning methods for quantitative radiomic biomarkers," *Scientific Reports*, vol. 5, pp. 1–11, 2015.
- [23] L. Senigagliaesi, M. Baldi, and E. Gambi, "Comparison of statistical and machine learning techniques for physical layer authentication," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1506–1521, 2021.
- [24] T. Wang, J. Deng, Y. She et al., "Radiomics signature predicts the recurrence-free survival in stage I non-small cell lung cancer," *The Annals of Thoracic Surgery*, vol. 109, no. 6, pp. 1741–1749, 2020.
- [25] N. Upadhyay and A. D. Waldman, "Conventional MRI evaluation of gliomas," *The British Journal of Radiology*, vol. 84, no. 2, pp. 107–111, 2011.
- [26] J. Wei, Z. Lu, K. Qiu, P. Li, and H. Sun, "Predicting drug risk level from adverse drug reactions using smote and machine learning approaches," *IEEE Access*, vol. 8, pp. 185761–185775, 2020.
- [27] Y. C. Zhang, A. Oikonomou, A. Wong et al., "Radiomics-based prognosis analysis for non-small cell lung cancer," *Scientific Reports*, vol. 7, pp. 1–8, 2017.
- [28] A. Golebiewska, A.-C. Hau, A. Oudin et al., "Patient-derived organoids and orthotopic xenografts of primary and recurrent gliomas represent relevant patient avatars for precision oncology," *Acta Neuropathologica*, vol. 140, no. 6, pp. 919–949, 2020.
- [29] E. I. Zacharaki, V. G. Kanas, and C. Davatzikos, "Investigating machine learning techniques for MRI-based classification of brain neoplasms," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 6, pp. 821–828, 2011.
- [30] J.-b. Qin, Z. Liu, H. Zhang et al., "Grading of gliomas by using radiomic features on multiple magnetic resonance imaging (MRI) sequences," *Medical Science Monitor*, vol. 23, no. 9, pp. 2168–2178, 2017.
- [31] D. Yang, G. Rao, J. Martinez, A. Veeraraghavan, and A. Rao, "Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma," *Medical Physics*, vol. 42, no. 11, pp. 6725–6735, 2015.
- [32] R. T. Shinohara, E. M. Sweeney, J. Goldsmith et al., "Statistical normalization techniques for magnetic resonance imaging," *NeuroImage: Clinical*, vol. 6, no. 8, pp. 9–19, 2014.

Research Article

A Distributed Congestion Control Strategy Using Harmonic Search Algorithm in Internet of Vehicles

Meiyu Pang ^{1,2}, Jianing Shen ^{1,2} and Lixiu Wu ^{1,2}

¹School of Internet of Things Engineering, Wuxi Taihu University, Wuxi, Jiangsu 214064, China

²Jiangsu Key Construction Laboratory of IoT Application Technology, Wuxi, Jiangsu 214064, China

Correspondence should be addressed to Meiyu Pang; pangmy@wxu.edu.cn

Received 22 January 2021; Revised 22 March 2021; Accepted 5 April 2021; Published 15 April 2021

Academic Editor: Chenxi Huang

Copyright © 2021 Meiyu Pang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the diversified requirements of network application QoS (Quality of Service) in the terminal equipment of Internet of Vehicles, this paper proposes a distributed congestion control strategy based on harmony search algorithm and the Throughput Evaluation Priority Adjustment Model (TEPAM) to ensure real-time transmission of high-priority data messages related to security applications. Firstly, the channel usage rate is periodically detected and the congestion is judged; then, in order to minimize delay and delay jitter as the goal, harmony search algorithm is utilized to perform global search to obtain a better solution for the transmission range and transmission rate. Secondly, packet priority and the TEPAM are applied to indicate the sending right of each packet. The data message priority and throughput percentage factor are used to express the transmission weight of each data message. Besides, the real-time evaluation of path state in MPTCP is carried out by the batch estimation theory model, which realizes the on-demand dynamic adjustment of the network congestion time window. Finally, SUMO, MOVE, and NS2 tools are used to create a VANET-like environment to evaluate the performance of the proposed congestion control strategy. Experimental results show that the proposed method is superior to other three methods in the four indicators of average delay time, average transmission rate, number of retransmissions, and packet loss rate compared with other advanced methods.

1. Introduction

The rise of intelligent transportation industry provides a safe and efficient traffic management system for traffic managers, drivers, and passengers. The vehicle interconnection network realizes the communication between vehicle-to-vehicle, vehicle-to-road, and vehicle-to-service centres by wireless communication technology, which realizes the organic combination of the human-vehicle-road environment and improves traffic safety efficiency. VANET (Vehicular Ad Hoc Networks) is a mobile ad hoc network composed of wireless communication between vehicles and between vehicles and base stations. Vehicles need to exchange information at any time in VANET. The communication between vehicles is usually divided into two types of safety information: Heartbeat Safety Message (HSM) and emergency alarm safety message. HSM is a message sent periodically to exchange information such as location, speed,

and driving direction of vehicles. When there is a dense vehicle scene, periodic HSM very easily causes channel congestion and affects driving safety. Besides, a large transmission range can enable safety information to reach longer distances in VANET, but it also increases transmission interference and channel competition. A high transmission rate can update information in time, but it can also cause channel saturation and increase the collision rate of information. Therefore, there is a need for a solution that can dynamically and reasonably adjust the transmission range and transmission rate to meet the delay and packet loss rate constraints. Due to the strong mobility of nodes and rapid topology changes in VANET, it is an NP-difficult problem to reasonably optimize the transmission range and transmission rate [1, 2]. Reference [3] reviewed the congestion control methods of VANET. We identified the relevant parameters and performance indicators that can be used to evaluate these methods and analysed each method

based on multiple factors, such as traffic type and active or passive and the mechanism to control congestion. Reference [4] proposed a control scheme based on Uniparted Objective Tabu Search (UOTS). It found the optimal value of transmission range and transmission rate by Tabu search according to the minimized delay function. However, this scheme did not consider the impact of the packet loss rate and number of retransmissions when calculating system delay. Thus, it cannot guarantee the solution quality, and convergence speed is slow.

This paper proposes a distributed congestion control strategy based on Harmony Search (HS) algorithm. While ensuring to meet the requirements of most network applications, it effectively improves the real-time performance of high-priority applications. The main innovations are as follows:

- (1) In many existing methods, it is difficult to ensure the quality of the solution and convergence speed without considering the problem of calculation delay and other issues that affect the packet loss rate and retransmission number. The proposed method uses HS to optimize the transmission distance and transmission rate and reduce the delay and delay jitter. It can achieve good performance in terms of delay, packet loss rate, and throughput.
- (2) Many existing methods are difficult to guarantee high-priority and real-time performance of sending data packets. The proposed method uses data message priority and TEPAM. The data message priority and the throughput percentage factor are used to express the weight of each data message; the batch estimation theory model is used to evaluate the path status in MPTCP in real time. It realizes the on-demand dynamic adjustment of the network congestion time window, effectively improving the throughput of high-priority applications on the network.

This paper is organized as follows: in Section 1, the research background and innovation of the proposed method are introduced. In Section 2, the related research in this field and the current research status are introduced. Section 2 describes the steps of the proposed method in detail; in Section 3, SUMO, MOVE, and NS2 tools are used to create a VANET-like environment to evaluate the performance of the proposed congestion control strategy. Section 4 summarizes the proposed methods and points out the future research direction.

2. Related Research

In recent years, with the frequent occurrence of traffic accidents, Intelligent Transportation Industry (ITS) based on VANET has emerged as the times require, providing a safe and efficient transportation system for drivers, passengers, and other traffic managers. Due to the rapid change of network topology of VANET, when vehicles are too dense, congestion will occur, which will cause packet loss and increase the transmission delay, so that it cannot provide

traffic information in time and threaten traffic safety. However, due to the shared wireless channel of VANET, it is challenging.

At present, the most proposed congestion control methods are mainly divided into the following categories: (1) reserving bandwidth for high-priority information by limiting the bandwidth occupation of the beacon frame; (2) dynamic Competition Window (CW) is used to adjust the transmission rate of packets; and (3) adjusting the transmission range by adjusting transmission power of the data packet. In [5], a distributed transmission rate control method DRCV was proposed to detect the local channel load periodically and estimate load in the next stage and to meet the requirements of channel load by dynamically adjusting the transmission packet rate. In [6], a distributed power adjustment algorithm D-FPAV was proposed to dynamically allocate transmission power for all vehicles to ensure that the network load is lower than the maximum beacon frame load.

However, the extra overhead of D-FPAV algorithm increases the channel load, and convergence speed of the algorithm is slow. In [7], a distributed congestion control algorithm based on Non-Cooperative Game Theory (NCGT) was proposed. The algorithm divided the vehicle into two groups, which represent the vehicle to participate in games, so that the channel load in the next stage does not exceed the total target load. This method improved the delivery rate and channel utilization to a certain extent.

Reference [8] proposed a control strategy based on a single-target Tabu Search algorithm in IoV (Internet of Vehicles). According to the minimized delay function, Tabu Search algorithm was applied to find the optimal value of transmission range and transmission rate. However, the influence of the packet loss rate and retransmission number was not considered when calculating the delay, so the quality of the solution cannot be guaranteed and the convergence speed was slow. Reference [9] introduced an MPTCP specifically used for multimedia applications in IoV. Reference [10] used MPTCP protocol to improve the efficiency of file transfer in IoV. Nevertheless, the traditional congestion control algorithm in Transport Control Protocol (TCP) is not appropriate for direct application in MPTCP protocol due to the multipath nature of MPTCP. Reference [11] proposed a distributed power adjustment algorithm to dynamically allocate transmission power for all vehicles and ensured that the network load was larger than the maximum beacon frame load. However, the additional cost of this algorithm increased the channel load, and the algorithm converges slowly. An average Packet Loss Rate (PLR) analysis model for VANET was proposed in reference [12]. An explicit expression for the upper bound of average PLR was obtained by using Taylor series expansion, Holder's inequality, and the relay probability relaxation, which can facilitate the selection of parameters at physical and MAC layers for a better PLR. Unfortunately, this method cannot ensure the real-time transmission of high-priority data packets. Therefore, the abovementioned methods can be further improved.

3. Proposed Method

Generally, vehicular communication is divided into two types of safety messages: HSM and emergency alarm safety message. Among them, HSM is a kind of periodic message, which is utilized to exchange information such as position, speed, and driving direction of vehicles. When a scene with dense vehicles appears, periodic HSM is very likely to cause channel congestion and affect traffic safety, as shown in Figure 1.

As shown in Figure 1, when using TCP to transmit data, there are usually three stages.

- (1) Start stage (0-knee): when the load on the network is small, with the increase in load the throughput on the network will increase approximately linearly.
- (2) Accumulation stage: the load on the network continues to increase. After reaching the upper limit of network capacity, the growth rate of throughput will gradually slow down.
- (3) Overload stage (cliff-∞): the load is continuously increased, so that the network is seriously congested, resulting in packet loss. At this time, the response time of the network increases significantly, the network tends to be paralyzed, and the throughput drops sharply.

3.1. Global Search Using Harmony Search. Congestion control includes two parts: congestion detection and congestion control. The congestion detection section detects congestion by periodically measuring channel utilization [13–16]. After congestion detection, the proposed congestion control algorithm is executed. Each node independently performs the proposed scheme and independently acquires the optimal value for its own transmission range and number. The flow chart of the dynamic distributed congestion control scheme is illustrated in Figure 2.

The main purpose of the VANET congestion control scheme is to reduce the number of the delays, jitter, packet loss, and retransmission. In the process of delay calculation, the packet loss rate and retransmission quantity are involved, which are representative. Therefore, this paper focuses on delay and delay jitter. Then, the harmonic Tabu search algorithm is used to optimize the transmission range and transmission rate of vehicle nodes according to the objective function. In VANET, packet transmission delay consists of processing delay (D_{proc}), queuing delay (D_{queue}), and transmission delay (D_{trans}). Processing delay is the time to execute various algorithms, queuing delay is the waiting time of packets in the queue to be transmitted, and transmission delay is the time required for propagation. The packet transmission delay is expressed as

$$\text{Delay} = (D_{\text{proc}} + D_{\text{queue}} + D_{\text{trans}}) \left[\frac{d}{\text{TX}} \right], \quad (1)$$

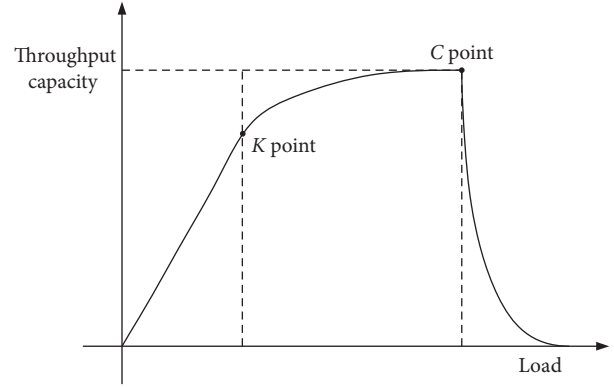


FIGURE 1: Channel congestion caused by dense vehicle environment.

where d is the distance between the sender and the receiver and TX is the transmission range. d/TX is the number of relay nodes between the sender and the receiver. Since the processing delay (ns level) is very small compared with other delays (ms level), it can be ignored. The queuing delay is calculated as follows:

$$D_{\text{queue}} = \frac{1}{\mu - \lambda} - \frac{1}{\mu} \cdot \frac{Q_L \rho^{Q_L}}{1 - \rho^{Q_L}}, \quad (2)$$

where ρ is the packet delivery rate, and its value is equal to λ/μ , where λ and μ are the number of successfully delivered packets and sent packets, respectively, and Q_L is the maximum queue length. The calculation formula of transmission delay is $D_{\text{trans}} = T_B + T_F$, in which T_B is the backoff delay and T_F is the frozen backoff delay. The fall back delay (T_B) is calculated as follows:

$$T_B = \begin{cases} \frac{W_{\min} \cdot \eta}{2} \cdot (2^{N_{RT}} - 1) & N_{RT} < m \\ \frac{\eta}{2} \cdot (W_{\max} - W_{\min} + W_{\max} \cdot (N_{RT} - m)) & N_{RT} > m \end{cases}, \quad (3)$$

where η is the length of the fall back delay slot, and its value is 20 us in VANET; W_{\max} is the size of the largest competition window, with a value of 1024 time slots; W_{\min} is the minimum competitive window size, with a value of 32 time slots; and m is the maximum number of fall back stages, with a value of $W_{\max} = 2^m W_{\min}$. According to IEEE 802.11, the value of m is 5; N_{RT} is the expected number of successful deliveries, which is calculated as follows:

$$N_{RT} = \sum_{s=1}^7 s P_c^{s-1} (1 - P_c) = \frac{1 - 8P_c^7 + 7P_c^8}{1 - P_c}, \quad (4)$$

where s is the number of fall back stages, and its value is 7 according to IEEE 802.11 standard; P_c is the collision probability, and its calculation is

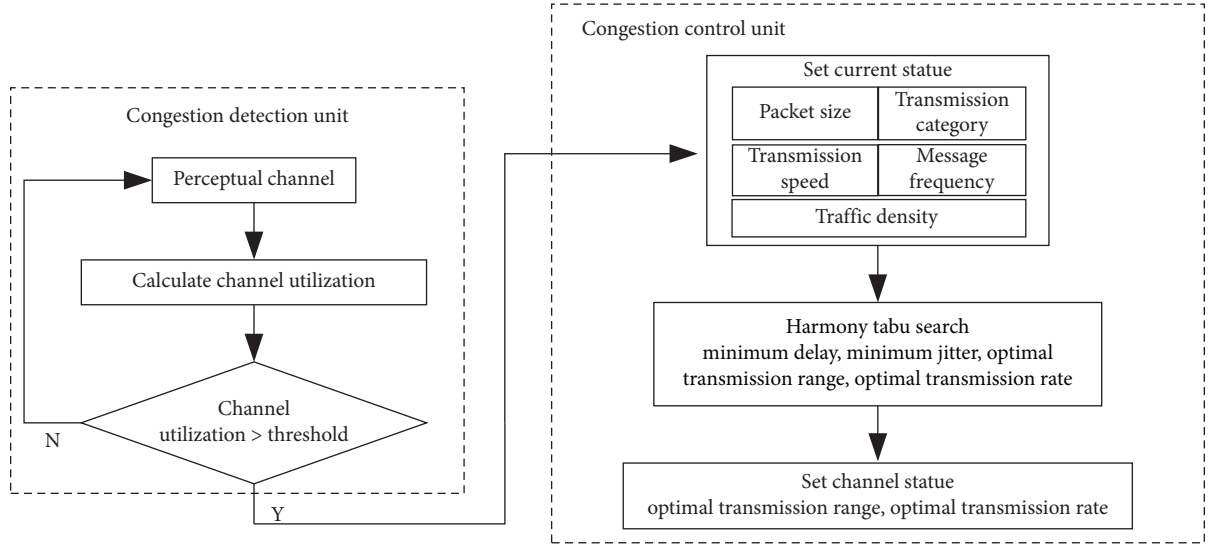


FIGURE 2: Flow chart of the congestion control scheme.

$$P_c = \frac{2W_{\min} \cdot N_c}{(W_{\min} + 1)^2 + 2W_{\min} \cdot N_c}, \quad (5)$$

where N_c refers to the number of competitors within the transmission range. When the channel is congested, the freeze back time T_F will appear. The calculation is as follows:

$$T_F = N_c + (N_c + 1) \cdot \left(\frac{N_{RT} - 1}{\sum_{t=2}^{N_c} t \cdot f(t)} \right) \cdot P_T, \quad (6)$$

where P_T is the time period of packet transmission including SIFs, DIFS, Data, and ACK; t is the number of competitors in the transmission range; and $f(t)$ is the number of packet retransmissions in case of collision.

$$f(t) = \binom{N_c}{t} \cdot \tau^t \cdot \left(\frac{N_c}{t!(N_c - 1)} \right) \cdot \tau^t \cdot (1 - \tau)^{N_c - 1}, \quad (7)$$

where τ is the transmission probability, which is calculated as $\tau = (2(1 - p_0))/(w_0 - 1)$, where w_0 is the current backoff window size, and its value is $8V$, V is the number of vehicles within the transmission range (V is the fixed value during propagation); and p_0 is the packet loss rate of MAC layer packets in each vehicle.

HS is a new heuristic search algorithm. In the performance of music, a wonderful harmony state can be achieved by repeatedly turning the notes of instruments in the band. HS algorithm is used primarily for fast search in Harmony Memory (HM). Therefore, HS algorithm can be utilized to optimize the transmission range and rate and minimize delay and delay jitter [17–20].

The main parameters of the HS algorithm are Harmony Memory Size (HMS), Harmony Memory Considering Probability (HMCR), and tuning probability (PAR), in which the HMS is a key parameter in HS. The larger the HMS is, the stronger the ability to locate a global optimal solution will be. However, excessive HMS, which is usually 5, will affect the speed of convergence to the optimal solution. HMCR has a great influence on the generation of the original solution, and a large value is conducive to the local search of the algorithm, which is 0.9. The tuning rate par, which controls local searches, allows the search to escape from local optimality with a value of generally 0.3.

Basic steps of HS are as follows:

- (1) HS parameters are initialized to set HMS, HMCR, and PAR parameters.
- (2) HM bank is initialized, and the initial matrix HM is composed of randomly generated HMS initial solutions.

$$HM = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & L & x_{N-1}^{(1)} & x_N^{(1)} \\ x_1^{(2)} & x_2^{(2)} & L & x_{N-1}^{(2)} & x_N^{(2)} \\ M & M & M & M & M \\ x_1^{(HMS-1)} & x_2^{(HMS-1)} & L & x_{N-1}^{(HMS-1)} & x_N^{(HMS-1)} \\ x_1^{(HMS)} & x_2^{(HMS)} & L & x_{N-1}^{(HMS)} & x_N^{(HMS)} \end{bmatrix}. \quad (8)$$

- (3) New solutions are generated by HM search, random selection, local disturbance, and other operations. The generation rules are as follows:

$$x_1^{(\text{new})} \leftarrow \begin{cases} \text{ran}\{x_1^{(1)}, x_1^{(2)}, L, x_1^{(HMS-1)}, x_1^{(HMS)}\}, & \text{ran}(1) \leq \text{HMCR} \\ \text{ran}(D), & \text{ran}(1) > \text{HMCR} \end{cases}, \quad (9)$$

where $\text{ran}(1)$ is a random number within the range of $[0, 1]$. If the search within HM produces the decision variable of new solution $x_i^{(\text{new})}$, then local perturbation of the new solution is required according to the probability par , and the rules are as follows:

$$x_i^{(\text{new})} \leftarrow \begin{cases} \text{Adjust } x_i^{(\text{new})} & \text{ran}(1) \leq \text{PAR} \\ \text{No Adjust } x_i^{(\text{new})} & \text{ran}(1) > \text{PAR} \end{cases}. \quad (10)$$

If $\text{ran}(1) \leq \text{PAR}$, then the decision variable is selected among its neighbours; otherwise, the original variable remains the same. The variable obtained after limited disturbance is $x_i^{(\text{new})} \leftarrow x_i^{(k=m)}$, ($m = 1, 2$).

- (4) HM is updated, the innovative solution is evaluated according to the fitness function, and it is replaced if it is better than the optimal solution in current HM.

3.2. MPTCP Link Condition Analysis. The abovementioned procedure is iterated, and HS is ended when a predetermined number of iterations or termination conditions are reached.

Link throughput varies with link load, and link conditions of MPTCP are predicted by analysing the proportion of throughput growth. Based on the proposed TCP link throughput model, throughput $x_r(t)$ at time t can be obtained by the following formula:

$$x_r(t) = \frac{1}{R_r(t)} \sqrt{\frac{2}{p_r(t)}}. \quad (11)$$

Here, $R_r(t)$ represents the delayed response time at time t on the path r , Round-Trip Time (RTT), and $p_r(t)$ represents the grouping loss rate at time t on the path $r \in M_u$, M_u is the set of all available paths for any application u . According to formula (11), the derivative of $x_r(t)$ can be obtained:

$$\frac{d(x_r(t))}{dt} = -\frac{\sqrt{2}}{2} \left(\frac{2}{R_r(t)^2 \sqrt{p_r(t)}} \frac{d(R_r(t))}{dt} + \frac{1}{R_r(t) p_r(t) \sqrt{p_r(t)}} \frac{d(p_r(t))}{dt} \right). \quad (12)$$

According to the analysis of formula (12), the throughput on path r at any time has the following $x_r(t)$ attribute:

- (1) If the delay response time $R_r(t)$ and packet loss rate $p_r(t)$ on path r at time t both increase, then $x_r(t)$ will decrease
- (2) If the delay response time $R_r(t)$ and packet loss rate $p_r(t)$ on path r at time t both decrease, then $x_r(t)$ will increase

Thus, the set C_r is determined to represent the changing trend of throughput for the MPTCP link.

$$C_r = \{t | t \in \arg(2p_r(t)\Delta R_r(t) + R_r(t)\Delta p_r(t) > 0)\}, \quad (13)$$

where C_r is derived from equation (12), which is the set of path r with gradually decreasing throughput at time t . $\Delta R_r(t)$ and $\Delta p_r(t)$ are for the convenience of calculation. $R_r(t) - R_r(t-1)$ and $p_r(t) - p_r(t-1)$ can be used to simplify the representation of p_r .

3.3. Throughput Evaluation Priority Adjustment Model.

The proposed method uses a batch estimation algorithm for throughput prediction, which is robust in one-dimensional nonlinear systems. The working process of the algorithm consists of two steps: the first step is adaptive, and the coefficient is calculated by batch estimation algorithm with measuring and calculating the data at time t and time $t-1$. The second step is prediction. The data from the first step are used to estimate the possible value at the moment $t+1$ by the algorithm. The estimated value obtained by batch

estimation algorithm represents the predicted throughput, which can predict the throughput of data stream in the future. $t+1$ hour throughput forecast $\hat{X}(t+1)$ for $t+1$ can be obtained from the following formula:

$$x(t+1) = \alpha(t)x(t) + (1 - \alpha_r(t)) \frac{x_r(t) + x_r(t-1)}{2}, \quad (14)$$

where the $\hat{X}(t+1)$ estimate for $t+1$ hour x , said on the estimate of alpha t time, estimates of the alpha available type (12) is obtained:

$$\alpha(t) = \frac{(2\alpha_r(t)^2)}{(\alpha_r(t)^2 + \alpha_r(t-1)^2)}. \quad (15)$$

$$\text{Here, } \alpha_r(t) = \sqrt{|x_r(t)^2 - \hat{x}(t)^2|}, \alpha_r(t-1) = \sqrt{|x_r(t-1)^2 - \hat{x}(t-1)^2|}.$$

Because of the recursive nature of algorithm, no additional information is mandatory, and the real-time prediction can be made using only the measured and estimated throughput values at time t and time $t-1$. The prediction method is simple to compute, and the robustness of algorithm can ensure the accuracy of predicted value as much as possible.

The dynamic priority of datagram is utilized to represent the real-time demand of the current datagram. It is described as follows:

$$\text{TP} = p + \left(\frac{R_r}{R_{\max}} \right) \left(\frac{\text{data}}{\text{data} + \alpha} \right), \quad (16)$$

where p is the priority of application. Depending on the general task scheduling of the operating system, the lower the p value is, the higher the priority will be. R is the

transmission path selected by application, R_r is the delay response time of the receiver when it was transmitted last time, R_{\max} is the maximum delay response tolerance time constant, after which the message is considered lost, the total data transmission amount of the current task of application is data, and α is the proportional control factor. According to the definition of the IEEE TCP protocol, the maximum data amount of a single data message is 1500 bytes. Therefore, $\alpha = 1500$ is chosen as the proportional control factor in this paper. That is, if the transmission data of an application is lower than the maximum data transmitted by a single data message, the transmission data of the application is relatively small.

The throughput ratio factor is utilized to represent the weight of network resources required by an application. The expression is as follows:

$$\beta = \left(\frac{TP - MID}{(N - M)/2} \right). \quad (17)$$

If the dynamic priority of an application is TP , the priority value range of the system is M to N . It can be seen from the definition that when the dynamic priority of transmission application is low, i.e., $TP \geq MID$, its throughput ratio factor $\beta \in [-1, 0]$ has no effect on guaranteeing the throughput occupation ratio of low-priority application. When the dynamic priority of transmission application is high, namely, $TP < MID$, its throughput ratio factor $\beta \in [0, 1]$, and the minimum throughput occupation ratio of high-priority applications in the current link can be properly guaranteed.

By means of throughput prediction, the congestion condition of the path can be obtained more accurately, which further improves the system transmission efficiency. However, the calculation amount is very small compared with the contemporary computer performance, so its impact on the transmission performance can be almost ignored.

On the other hand, the proposed method mainly considers the transmission requirements of high-priority tasks, so it will enhance the transmission time of low-priority tasks while reducing the transmission time of high-priority tasks. At the same time, due to the limitation of algorithm packet loss policy, when the transmission processes in the system are all extremely high-priority processes, or all extremely low-priority processes, the system transmission efficiency will be greatly affected, but this problem can be avoided by applying settings.

4. Experimental Results and Analysis

4.1. Experimental Parameter Settings. In the experiment, the network simulator NS2 was used to build the traffic system model, and the vehicle-mounted network mobile model generator was used to connect the traffic simulation tool SUMO with the network simulation tool NS2. Therefore, with SUMO, MOVE, and NS2 tools, a VANET-like environment can be created to evaluate the performance of proposed congestion control strategy. The experiment is considered as an urban traffic scenario (4 lanes, 2 in each

direction), in which communications between vehicles and the base station follows the IEEE 802.11p protocols [21, 22].

In addition, data transmission in the MAC layer has carrier monitoring multiple access mechanisms, and the propagation model of wireless signal uses the Nakagami attenuation model. The simulation time is configured on 200 s, where the transmission range and transmission rate are 10~1000 m and 3~27 Mbps, respectively. The transmission range values are 10, 50, 100, 125, 150, 210, 300, 350, 450, 550, 650, 750, 850, 850, 930, 970, and 1000 m. The standard values set for transmission rates are 3, 4.5, 6, 9, 12, 18, 24, and 27 Mbps. So, there are 136 combinations of the transmission range and the transmission rate. Parameters of urban traffic scenes are shown in Table 1.

4.2. Comparison and Analysis of Results. In order to evaluate the performance of the proposed congestion control strategy, the proposed method was compared with the algorithms in [8, 9, 12], and the following four indicators were set:

- (1) The average delay time of data packet transmission from the sender to receiver
- (2) Mediocre throughput of the average rate at which information is successfully transmitted in a transmission channel
- (3) Packet loss rates of the ratio of the number of lost packets to the number of transmitted packets
- (4) The number of retransmitted packets (each experiment was performed 20 times and the average value was estimated)

4.2.1. Results Varied with the Number of Vehicles. The average delays varied with the number of vehicles are shown in Figure 3. The throughput varied with the number of vehicles is shown in Figure 4. The packet loss rate varied with the number of vehicles is shown in Figure 5. The retransmission quantity varied with the number of vehicles is shown in Figure 6.

As can be seen from Figure 3, with the increase of the number of vehicles, the possibility of channel congestion becomes larger, resulting in increased delay. Surprisingly, the proposed method gets the smallest delay. When the number of vehicles is 200, the proposed method reduces the delay by about 76.8%, 73.6%, and 33.6%, respectively, compared with [8], [9], and [12].

As can be seen from Figure 4, an increase in the number of vehicles results in an increase in the number of communication packets, which increases throughput. When the number of vehicles is 200, the average throughput in [8], [9], and [12] is 17.55, 22.23, and 12.38 Mbps, respectively. Since the reduction of channel collisions and the improvement of channel utilization can be effectively improved by the proposed method, the highest network throughput is achieved.

As shown in Figure 5, the packet loss rate increases with the number of vehicles. Because the more the vehicles are

TABLE 1: Parameters of urban traffic scenes.

Parameter	Value (city)
Scene: the size	1500 m × 1000 m
Number of lanes	4 (2 in each direction)
Number of vehicles	60, 80, 110, 140, 180, 190
The speed of the car	0–60 km/h
Transmission range	10–1000 m
Transmission rate	3–26.8 Mbps
Run number	20
Competitive window size	15–1023
The band width	10 MHz
Safety information	10 package/s
MAC model	802.11p
Signal propagation model	Nakagami ($m = 3$)
The simulation time	200 s

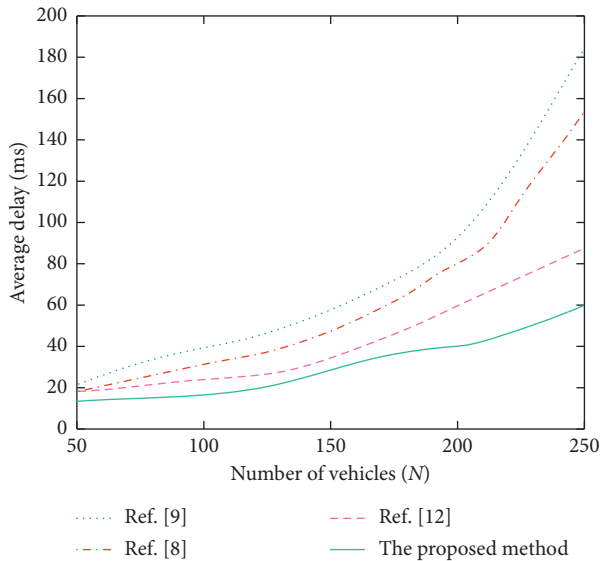


FIGURE 3: Average delays varied with the number of vehicles.

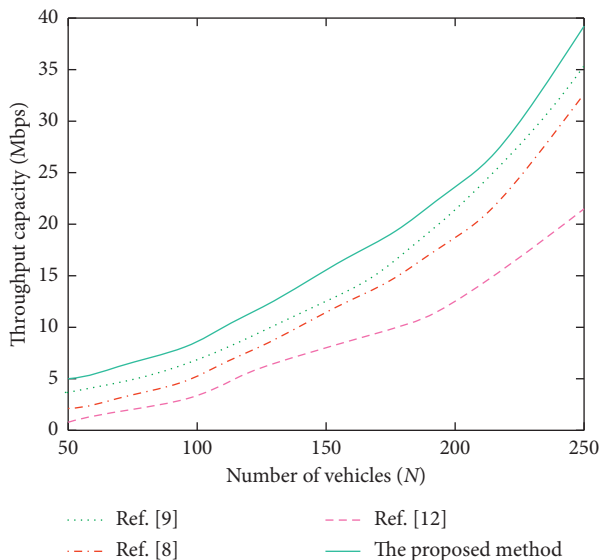


FIGURE 4: Throughput varied with the number of vehicles.

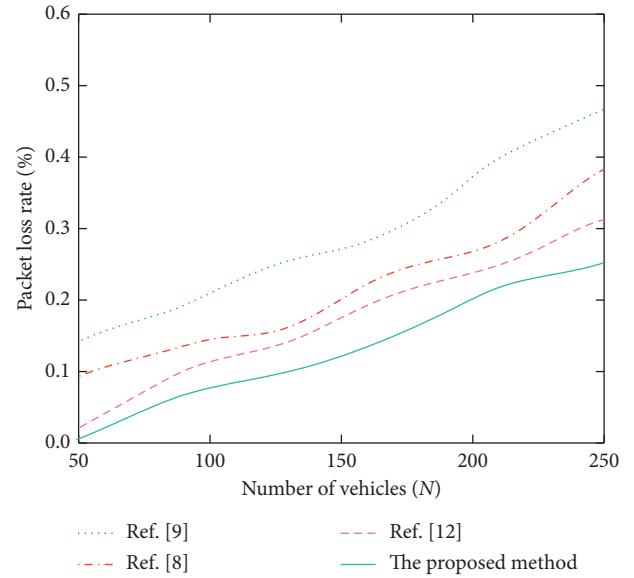


FIGURE 5: Packet loss rate varied with the number of vehicles.

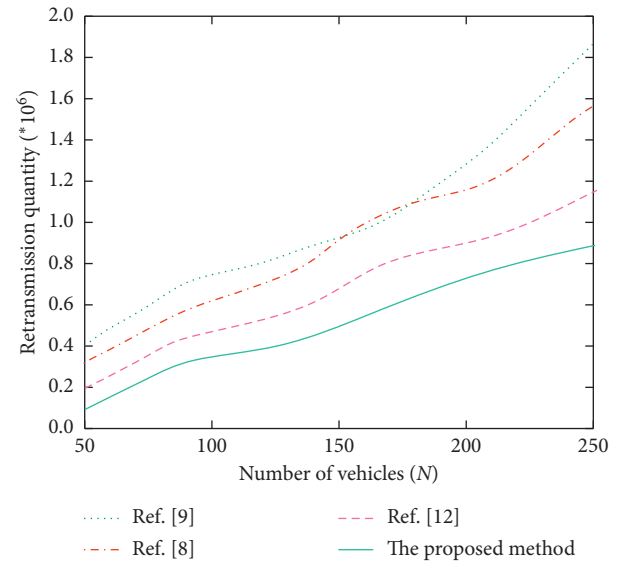


FIGURE 6: Retransmission quantity varied with the number of vehicles.

there, the greater the communication pressure and interference of the channel will be. However, the proposed method has the lowest packet loss rate. Compared with algorithms in [8], [9], and [12], when the number of vehicles is 200, the proposed method has reduced the packet loss rate by 37.8%, 28.8%, and 12.9%, respectively. When packets are lost, they need to be retransmitted, so a lower packet loss rate can decrease the number of retransmitted packets with transmission delay and then increase the throughput. The packet loss rate of algorithms in [8] and [12] fluctuates greatly in the transmission process. The algorithm in [9] has a high packet loss rate due to excessive link congestion. The adjustment effect of congestion avoidance algorithm on the congestion window value is more intuitively shown.

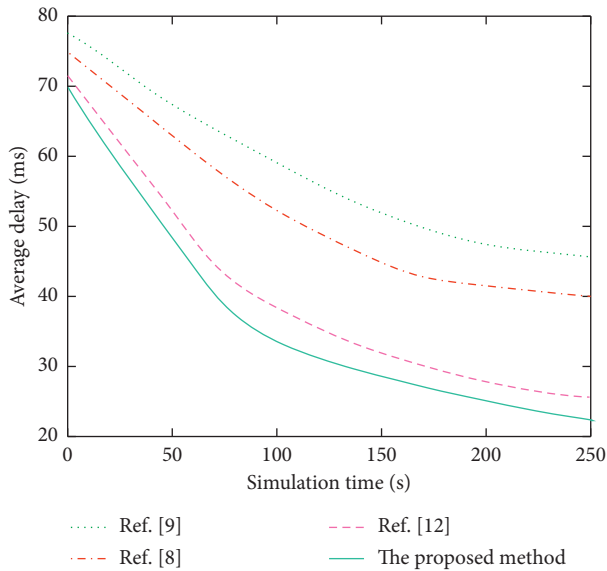


FIGURE 7: Average delay of the simulation time.

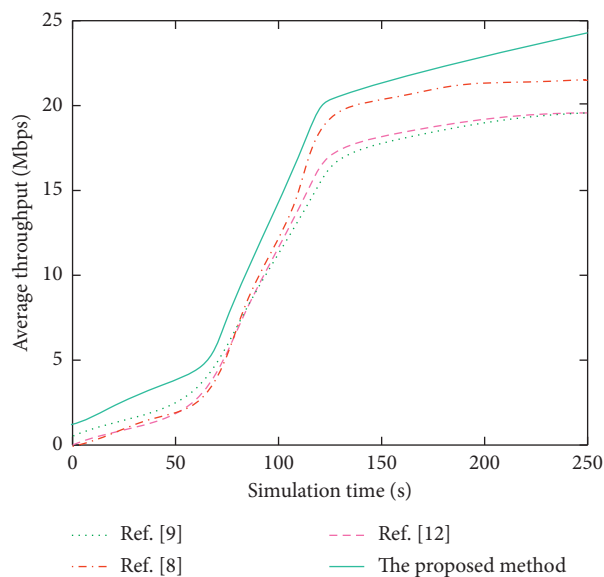


FIGURE 8: Average throughput with the simulation time.

As shown in Figure 6, the data packet retransmission numbers can be reduced by using the proposed method. When the number of vehicles is 200, compared with methods proposed in [8], [9], and [12], it can be reduced by 46.2%, 36.8%, and 22.8%, respectively.

4.2.2. Results Varied with the Simulation Time. The average delay and throughput varied with the simulation time are shown in Figures 7 and 8, respectively, and the simulation time was set as 50, 100, 150, 200, and 250 s, respectively.

As is shown in Figure 7, at the beginning of the simulation, the higher collision in transmission leads to higher delay, but it will decrease rapidly with the operation of various congestion control algorithms. From the 50 s to the 250 s, the average delay in [8], [9], and [12] and the proposed

method is decreased by 36.8%, 42.6%, 58.8%, and 62.8%, respectively. Among them, the proposed method owns the lowest average delay.

As shown in Figure 8, at the beginning of the simulation, the throughput of various algorithms is similar, and with the execution of the optimization process, the throughput increases obviously until reaching saturation. The throughput of the proposed method is the highest because the proposed method automatically adjusts the transmission range and transmission rate of vehicle nodes and effectively improves the throughput.

5. Conclusions

Based on the existing congestion control algorithm MPTCP, this paper proposes a distributed congestion control strategy based on HS and TEPAM according to the data transmission requirements that will appear in the future IoV system. The proposed congestion control algorithm can solve the real-time data transmission problem with high-priority requirements and further improve the data transmission efficiency of MPTCP. Simulation results show that the proposed algorithm significantly improves the data transmission speed of high-priority applications compared with other congestion control algorithms. Furthermore, it improves the overall data transmission efficiency, which helps improve traffic safety.

Although the congestion control algorithm proposed in this paper has been relatively perfect, it cannot meet the priority transmission requirements of security applications in the Internet of Vehicles and other systems. The future research direction is to determine the priority based on the importance of information and further rationally adjust the transmission method to ensure the priority and safe transmission of important information. In addition, we will consider appropriately to adjust the packet loss strategy of algorithm according to the priority of current transmission task in system, as well as adjust the throughput prediction model to better adapt to various network environments.

Abbreviations

MPTCP:	Multipath transport control protocol
TCP:	Transport control protocol
DRCV:	Distributed transmission rate control method
HSM:	Heartbeat safety message
HM:	Harmony memory
HS:	Harmony search
HMS:	Harmony memory size
HMCR:	Harmony memory considering probability
PAR:	Tuning probability
TEPAM:	Throughput evaluation priority adjustment model.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Natural Science Foundation of China (No. 61873112).

References

- [1] J. Eze, S. Zhang, E. Liu, and E. Eze, "Cognitive radio-enabled Internet of Vehicles: a cooperative spectrum sensing and allocation for vehicular communication," *IET Networks*, vol. 7, no. 4, pp. 190–199, 2018.
- [2] M. Ayob, M. Z. A. Nazri, and N. R. Sabar, "Meta-harmony search algorithm for the vehicle routing problem with time windows," *Information Sciences*, vol. 325, pp. 140–158, 2015.
- [3] X. Liu and A. Jaekel, "Congestion control in V2V safety communication: problem, analysis, approaches," *Electronics*, vol. 8, no. 5, p. 540, 2019.
- [4] K. W. Jang, "A tabu search algorithm for routing optimization in mobile ad-hoc networks," *Telecommunication Systems*, vol. 51, no. 3, pp. 177–191, 2012.
- [5] M. Javed and J. Y. Khan, "Performance analysis of a time headway based rate control algorithm for VANET safety applications," in *Proceedings of the 2013 7th International Conference on IEEE Signal Processing and Communication Systems (ICSPCS)*, pp. 1–6, IEEE, Carrara, Italy, 2013.
- [6] M. Torrent-Moreno, J. Mittag, P. Santi et al., "Vehicle-to-vehicle communication: fair transmit power control for safety-critical information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3684–3703, 2011.
- [7] D. Wu, Y. Ling, H. Zhu et al., "The RSU access problem based on evolutionary game theory for VANET," *International Journal of Distributed Sensor Networks*, vol. 30, no. 3, pp. 276–277, 2013.
- [8] C.-J. Hsu, J. Fikentscher, and R. Kreeb, "Development of potential methods for testing congestion control algorithm implemented in vehicle-to-vehicle communications," *Traffic Injury Prevention*, vol. 18, no. 1, p. S51, 2017.
- [9] M. Ramezani and E. Ye, "Lane density optimisation of automated vehicles for highway congestion control," *Transportmetrica B: Transport Dynamics*, vol. 77, no. 5, pp. 1–21, 2019.
- [10] W. S. Jung, J. Yim, and Y. B. Ko, "Adaptive offloading with MPTCP for unmanned aerial vehicle surveillance system," *Annals of Telecommunications*, vol. 46, no. 11, pp. 1–14, 2018.
- [11] X. Hu, C. M. Martinez, and Y. Yang, "Charging, power management, and battery degradation mitigation in plug-in hybrid electric vehicles: a unified cost-optimal approach," *Mechanical Systems and Signal Processing*, vol. 87, no. 7, pp. 4–16, 2017.
- [12] W. Lai, N. Wei, W. Hua et al., "Analysis of average packet loss rate in multi-hop broadcast for VANETs," *IEEE Communications Letters*, vol. 22, no. 1, pp. 157–160, 2017.
- [13] L. D. Cicco, G. Carlucci, and S. Mascolo, "Congestion control for WebRTC: standardization status and open issues," *IEEE Communications Standards Magazine*, vol. 1, no. 2, pp. 22–27, 2017.
- [14] A. A. Ahmed and W. Ali, "A lightweight reliability mechanism proposed for datagram congestion control protocol over wireless multimedia sensor networks," *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 3, pp. 32–36, 2018.
- [15] W. Noh, W. Shin, T. Lee, and H.-H. Choi, "Distributed uplink interference control based on resource splitting in heterogeneous cellular networks," *Wireless Networks*, vol. 23, no. 2, pp. 625–640, 2017.
- [16] Y. Yukun, L. Jiangbing, X. Dongliang et al., "Centralized congestion control routing protocol based on multi-metrics for low power and lossy networks," *The Journal of China Universities of Posts and Telecommunications*, vol. 24, no. 5, pp. 39–47, 2017.
- [17] M. A. Jan, P. Nanda, X. He et al., "PASCCC: priority-based application specific congestion control clustering protocol," *Computer Networks*, vol. 74, no. 1, pp. 92–102, 2017.
- [18] C. W. Feng, L. F. Huang, C. Xu et al., "Congestion control scheme performance analysis based on nonlinear RED," *IEEE Systems Journal*, vol. 11, no. 4, pp. 224–225, 2017.
- [19] S. Mohanarangan and D. Sivakumar, "Congestion control through policy based scheduling algorithm," *Journal of Computational and Theoretical Nanoscience*, vol. 14, no. 9, pp. 315–319, 2017.
- [20] D. M. Chiu and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," *Computer Networks & ISDN Systems*, vol. 17, no. 1, pp. 1–14, 2018.
- [21] R. Sanchez-Iborra and M. D. Cano, "On the similarities between urban traffic management and communication networks: application of the random early detection algorithm for self-regulating intersections," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 4, pp. 48–61, 2017.
- [22] R. A. Mallah, A. Quintero, and B. Farooq, "Cooperative evaluation of the cause of urban traffic congestion via connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 1–9, 2019.

Research Article

A Single-Label Model to Ensure Data Consistency in Information Security

Cigdem Bakir 

Department of Computer Engineering, Yildiz Technical University, Istanbul 34220, Turkey

Correspondence should be addressed to Cigdem Bakir; cigdem.bakr@gmail.com

Received 5 March 2021; Revised 23 March 2021; Accepted 26 March 2021; Published 2 April 2021

Academic Editor: Yi-Zhang Jiang

Copyright © 2021 Cigdem Bakir. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Information security is defined as preventing actions such as unauthorized access and use, modification, and removal of information. It consists of certain basic elements of confidentiality, integrity, and accessibility. There are numerous studies in published literature which have been conducted to ensure information security. However, there is no previous study that covers these three basic elements together. In the present study, a model that includes these three key elements of information security together for big data was proposed and implemented. With this proposed “single-label model,” a more practical and flexible structure was established for all operations (read, write, update, and delete) performed on a database on real data. In previous studies conducted with a label model, separate labels were used for read-only or write-only operations, and there was no structure that could ensure both confidentiality and integrity at the same time. The present study, however, shows what type of authorization and access control could be established between which processes and which users by looking at a single label for all the operations performed on the data. Thus, in contrast to the previous studies seen in published literature, data confidentiality, data integrity, and data consistency were all guaranteed for all transactions. The results of the proposed single-label model were also shown comparatively by conducting an experimental study of its application. The results obtained are promising for further studies.

1. Introduction

Information security is defined as preventing actions such as unauthorized access and use, modification, and removal of information, and it consists of certain basic elements including confidentiality, integrity, and accessibility [1, 2]. Confidentiality is the protection of information against being accessed, read, or used by unauthorized persons in any way. Integrity is the prevention of modification of information by unauthorized persons and the preservation of its original nature. Accessibility, on the other hand, is that the information is accessible and readily available as long as it is needed.

Today, there are new and highly effective threats that damage information systems and resources [3]. Although there are many measures taken to protect systems from such harmful threats that are supported by advanced technologies, it has been seen that attackers can still often succeed. In

these and similar cases, any incident that causes a violation of any of the three basic elements of information security (confidentiality, integrity, and accessibility) is considered to be a security problem [4]. While some violations intentionally make systems inaccessible and disrupt services, others occur accidentally due to unforeseen faults. Whether accidental or malicious, security violations seriously affect the activity and reliability of an institution.

In general, threats often turn into attacks by exploiting gaps or vulnerabilities in systems. Therefore, it can be said that it is of great importance to provide all these three basic elements together to prevent such attacks from damaging information systems. In short, no matter how secure a system is, the important thing here is to ensure control of the access and authorization processes that may allow any attack [5].

Some leading factors that cause security breaches (or violations) include Denial of Service (DOS) attacks,

Distributed Denial of Service (DDOS) attacks, inappropriate web browsing behavior, wiretapping, access to resources using a backdoor, and data changes occurring accidentally or intentionally [6]. Data that is deliberately or accidentally changed directly affects the integrity principle of information systems security in particular, and it results in an emerging security breach. The occurrence of such data modification events, like giving excessive authorization to users and exercising poor control of permissions, plays an important role [7]. To deal with such problems, a model designed according to the specific access rights (e.g., read, write, update, and delete) is required for organizations and users. However, studies have shown that these models are unable to fully meet the needs of rapidly growing and increasingly complex systems, because they represent a serious financial burden and fail to fully provide information flow control [8–11]. Therefore, it is seen that it is not enough for information systems to be constructed in a way to protect them only from unauthorized access, malicious users, and misuse. In this study, a model was created to provide the three basic elements of information security together by using real data. In this way, no user or group of users would be able to access data that is not authorized at their level or data that they are not allowed to perform various operations on.

In this study, a single-label model is created. The scientific contribution of this model is that while the data available to be used by the stakeholders can only easily be used by authorized actors, it does not allow the use of these data by unauthorized third-party actors. At the same time, this model contributes to the research of methods that enable the use of jointly used resources without causing information leakage. Therefore, in this study, we describe a distributed label model that can maintain data confidentiality with information flow control in distributed databases. The difference between this study and the other studies on this subject is that this label model targets data confidentiality and integrity among nonreliable actors and environments. Through the labels given to the data, each actor can determine his/her own security policy independently from other actors and authorize the ones that he/she chooses. The purpose of this study was to develop a method that allows different users to access the data in a distributed environment and protects confidentiality. It was aimed at investigating methods preventing unauthorized access to data being accessed jointly by multiple actors.

In the remainder of this study, other researches related to this subject are presented in Section 2, while the method is presented in Section 3. The proposed model is discussed in Section 3, and its application is detailed in Section 4. Section 5 details the evaluation and conclusions.

2. Related Works

Information is a valuable asset. Therefore, access, processing, updating, deleting, and authorizing operations should be carefully managed to ensure that confidentiality, integrity, and accessibility are maintained. In recent years, some techniques have been developed in published literature

which outline the rules related to access, authorization, monitoring, and control of information and information systems [12–14]. However, it is seen in many industries that the development area of these techniques has narrowed and that existing techniques do not fully meet the new business requirements that arise with developing technology, and they cannot be managed in accordance with the organizational structure. In addition, serious costs arise in the progress towards a manageable model, and the dynamism that is necessary for the use and sharing of resources is not achieved.

In recent years, various studies using different techniques for the purposes mentioned have been described in published literature. Schultz and colleagues developed a platform that allowed the data access of users to be automatically tracked. Because a user logs into the system separately for each transaction, authority control is performed again. The user has to perform the authority check at each stage. If he/she does not perform the check at any one stage, data confidentiality is breached. This creates the need for automatic monitoring of authority [15]. Parker et al. presented a platform extension for database transactions. In this platform, each table has a label and protects its length [16], but this method can impose high computational costs and high overheads. Yang et al. used information flow control in web applications, but this approach can be expensive in both space and time and requires more memory [17]. Muthukumaran et al. applied information flow control (IFC) with FlowWatcher monitoring software that provides applications with a web proxy but limits the granularity of policies it can enforce [18]. In previous studies in published literature [19–22], a separate label was used for each operation (read, write) carried out on the object, and only reading and writing were performed. In the present study's proposal, by contrast, all operations performed on the object (read, write, update, and delete) are carried out using a single label. In this way, by looking at a single label, what type of authorization style is used between which operations and which actors can be understood.

In recent years, various studies using different techniques for the purposes mentioned above have been described in published literature [13, 15, 23–26]. In this present study, on the other hand, there is no need for separate control for both authorizing and denying authorization. There is no need for separate authorization or access control for each operation such as reading, writing, updating, and deleting. In addition, by tracking the access of malicious actors to data, attempts are made to prevent information disclosure.

Fog computing or fog networking, also known as fogging, is pushing the frontiers of computing applications, data, and services away from a centralized cloud to a logical stream on the network edge. Fog networking systems work on building the control, configuration, and management over the Internet backbone [27].

Software-defined networking (SDN) is a promising approach to networking which provides an abstraction layer for the physical network [28]. In published literature, a recurrent neural network (RNN) model based on a new

regularization technique (RNN-SDR) was proposed by the authors. This technique supported intrusion detection within SDNs [28]. Nevertheless, this model is not practical for implementation in the context of an SDN. Prete and Schweitzer contextualized the existing problems in current computer networks and presented the SDN network as one of the main proposals for the viability of the Internet of the future. Simulations were created in an SDN network scenario using a POX Controller [29]. However, there is a need to obtain a synergistic effect that will make cloud environments more efficient, dynamic, and flexible, including automatic reconfiguration of network clusters.

In the current study, a single-label model was developed. The scientific contribution of this model is that while the data available to use by the stakeholders can be used easily only by authorized actors, it does not allow the use of these data by unauthorized third-party actors. At the same time, this model contributes to research into methods that enable the use of jointly used resources without causing information leakage. Therefore, in this study, a single-label model was developed which can maintain data confidentiality and integrity with information flow control. The difference between this study and other studies with a single-label model is that it targets data confidentiality and integrity of users. Through the labels given to the data, each actor can determine his/her own security policy independently from the other actors and authorize the ones that he/she chooses from the other actors. Moreover, access control and authorization are ensured in accordance with the actor's wishes, without causing data leakage and with the supervision of information flow control. The actors are able to create their own security, confidentiality, and integrity policies in a practical and flexible way. The difference between this study and other studies is that it provides data confidentiality, data integrity, and data consistency together.

3. Proposed Model

The single-label model consists of actors, objects, and labels.

3.1. Actor. The actors include data owners and users or groups of users who perform operations such as granting and receiving data authorization. Each actor labels his/her data for data confidentiality and integrity. The label consists of a list of security policies that are provided by the actors. Each actor labels his/her data for data privacy. That is, a label is determined which is paired with a data object. In addition, each actor has the right to safely change these security policies separately. Figure 1 shows a sample actor hierarchy. In this figure, X and Y are the representatives of a worker group. Worker Z has two tasks and duties as an engineer and a unit head. In the principal hierarchy, the process of granting authority is transitive. For instance, $X \rightarrow Y$ stands for granting authority by X to the principal Y. If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$ is also true.

3.2. Label. A label is a collection of policies that are created for the protection of data. That is, a label is determined

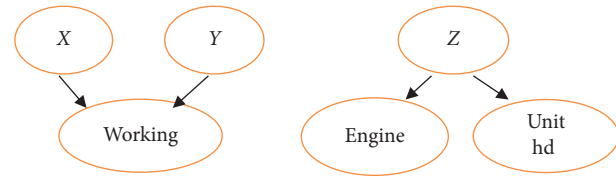


FIGURE 1: Examples of the principal hierarchy.

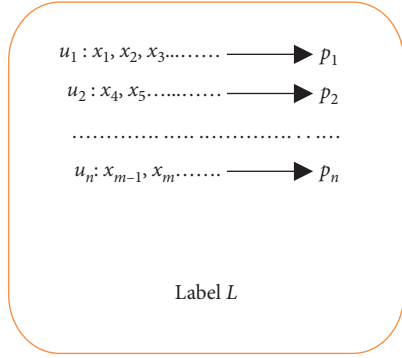
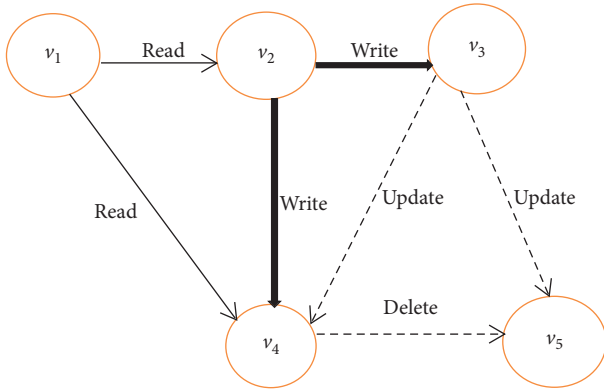
which is paired with a data object. In addition, each actor has the right to safely change these security policies separately. This model was developed for unreliable actors and environments. All actors change their own policy independently of each other. The object consists of data to which authorization is granted or received by actors. The label consists of the list of security policies issued by actors. Each actor labels his/her data for data confidentiality. In addition, each actor separately has the authority to safely change these security policies.

Figure 2 shows the contents of a label. Here, while u_1, u_2, \dots, u_n show the owners of the data object from the actors in the system, the terms x_1, x_2, \dots, x_m refer to the actors to whom authorization is given for any transaction by the data owners: p_1, p_2, \dots, p_m , that is, each content definition on the L label, shows the security policy of the relevant actor regarding these common data. Each actor who owns a data object determines his/her own policy on the label. Then, one of the actors sends these data objects to the other actors with its label.

3.3. Graph Modeling of Labels. In previous studies in published literature [19–22, 30–32], a separate label has been used for each operation (read, write) carried out on the object, and only reading and writing have been performed. However, in the present study proposal, all operations performed on the object (read, write, update, and delete) are carried out using a single label. In this way, by looking at a single label, what type of authorization style there is between which operations and which actors is understood.

In this present study, the single-label model is shown by a graph data structure (Figure 3) in which we let the label determined for graph G be L_G . In this study, the circles in the graph data structure show the actors. Which operation will be performed in the distributed database is determined by the way the arrow is drawn. A different arrow is used for each of the read, write, update, and delete operations. Thus, with a single label, a more practical and more secure authorization and access operation is created.

L_G consists of five parts, namely, owner, readers, writers, updaters, and deleters. The way the arrows are drawn in the graph show the types of authority needed to access the data. Here, while “owner” denotes the actors who own the labeled object, “readers” refers to the actors to whom authorization is given to read data owners’ transactions; “writers” refers to the actors to whom authorization is given to write to the data owners’ transactions; “updaters” refers to the actors to whom authorization is given to update the data owners’ transactions; and “deleters” refers to the actors to whom

FIGURE 2: Label L example for the data object.FIGURE 3: A graph G modeling of the label (L_G).

authorization is given to delete data owners' transactions. The label shown in Figure 1 combined with graph G can be expressed in the L_G typing format as follows:

$$L_G = \{v_1 : v_2, v_4; v_2 : v_3, v_4; v_3 : v_4, v_5; v_4 : v_5; v_5\}. \quad (1)$$

The semicolon used when creating a label separates the policies from one another. Accordingly, the L_G label has five policies: $\{v_1:v_2, v_4\}$, $\{v_2:v_3, v_4\}$, $\{v_3:v_4, v_5\}$, $\{v_4 : v_5\}$, and $\{v_5 : \}$. While v_1, v_2, v_3 , and v_4 denote the owners of the data object to which the L_G label belongs, v_2, v_3, v_4 , and v_5 represent the actors authorized by the data owners for various object transactions (read, write, update, and delete).

Let us assume that the first policy shows the read operation on the object.

The first policy is expressed with the $v_1 \rightarrow v_1, v_1 \rightarrow v_2$, and $v_1 \rightarrow v_4$ edges. This means that the v_1 actor allows the v_1, v_2 , and v_4 actors to read his/her data.

Let us assume that the second policy shows the write operation on the object.

The second policy is expressed with the $v_2 \rightarrow v_2, v_2 \rightarrow v_3$, and $v_2 \rightarrow v_4$ edges. This means that the v_2 actor allows the v_2, v_3 , and v_4 actors to write to his/her data.

Let us assume that the third policy shows the update operation on the object.

The third policy is expressed with the $v_3 \rightarrow v_3, v_3 \rightarrow v_4$, and $v_3 \rightarrow v_5$ edges. This means that the v_3 actor allows the v_3, v_4 , and v_5 actors to read his/her data.

Let us assume that the fourth policy shows the delete operation on the object.

It is expressed by $v_4 \rightarrow v_4$ and $v_4 \rightarrow v_5$ edges. This means that the v_4 actor allows the v_4 and v_5 actors to delete his/her data.

The last policy is expressed with the $v_5 \rightarrow v_5$ edge. This means that v_5 does not allow anyone other than himself/herself to perform any transaction on his/her data.

3.4. Bank Example. A bank has many customers. Each bank is obliged to protect and save its customers' account information such as money, goods, and investments from other customers or noncustomer principals. In Figure 4, a bank's customer operations have been shown by employing label modeling. In this figure, the oval shapes are as follows: M is customer, B is bank, and T is the principal's computing customer assets. Arrows represent information flow between principals, while squares represent the database and the data.

Any customer can, by labeling i ($1 \leq i \leq n$) assets with $\{M_i:B, M_i\}$, forge their own security policy. Also, each customer performs operations such as drawing or depositing cash and so forth at different times. A bank has to conduct these operations safely. These banks label all customer operations performed with $\{M:B, M\}$. Thus, banks can read customers' information. Customer i operations, like withdrawing cash, depositing cash, money transfers, and so forth, are conducted by the T principal. T is a program computing customers' asset details. The T principal can declassify any asset information that each i customer labels with $\{M_i:B, M_i\}$, and with a $\{B:B\}$ label it transfers them to the bank's database. Thus, this bank can control the flow of information and, to ensure that other principals in the system cannot read these data, it saves these data with a $\{B:B\}$ label in its private database. These labels are created for all operations performed in the database and combine them into one label.

4. Experimental Study

When the proposed single-label model was compared with the double-label model in published literature, and the performance results obtained in terms of accuracy and time are given in the following sections.

4.1. Accuracy. In Table 1, the success of the proposed single-label model and that of the double-label model in published literature are compared against a real data set, which has been taken from a hospital and whose classes are obvious. Accuracy rates were calculated for about 100 actors and 20 objects randomly selected from this data set. In addition, all classes of this data set were specified. Accuracy rates were calculated according to their real class. While measuring the accuracy rate, the classes of the model created for this study were calculated by comparing them with real classes. The success of the proposed model is clearly shown in Figure 5. When the performances of both methods were compared for all operations performed on objects in terms of accuracy

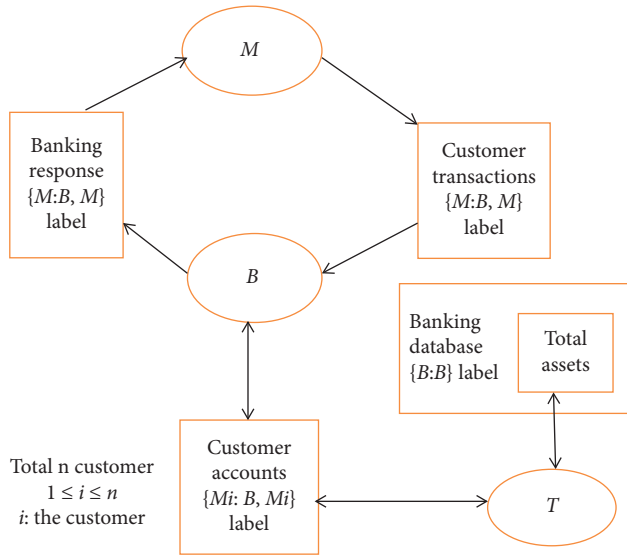


FIGURE 4: Labels for banking accounts.

TABLE 1: Accuracy rates for 100 actors and 20 objects.

Accuracy rate (%)	Double label	Single label
Read	87.27	99.94
Write	89.17	98.61
Update	79.50	96.37
Delete	83.34	97.81

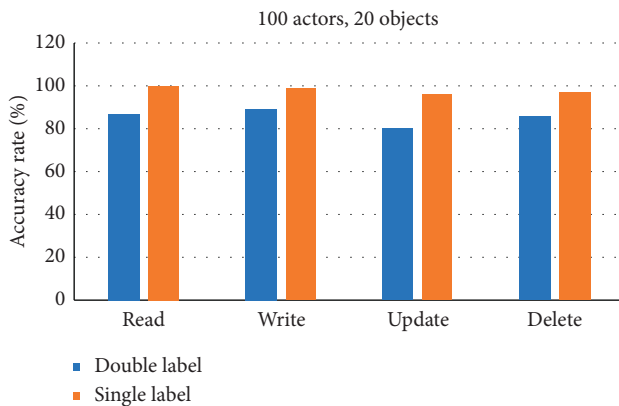


FIGURE 5: Accuracy rates for 100 actors and 20 objects.

rates, the success of the proposed single-label model can be clearly seen. In particular, it gives more successful results in reading and deleting operations. This is because writing and updating operations are more difficult than other operations.

In Table 2, the success of the proposed model (single label) and that of the model in published literature are compared in terms of accuracy. Accuracy rates have been calculated for about 1000 actors and 200 objects. The success of the proposed model is clearly shown in Figure 6. When the performances of both methods are compared in terms of accuracy rates for all operations performed on objects, the success of the proposed model can be clearly seen.

TABLE 2: Accuracy rates for 1000 actors and 200 objects.

Accuracy rate (%)	Double label	Single label
Read	84.21	96.93
Write	85.17	95.58
Update	82.96	91.58
Delete	81.55	95.10

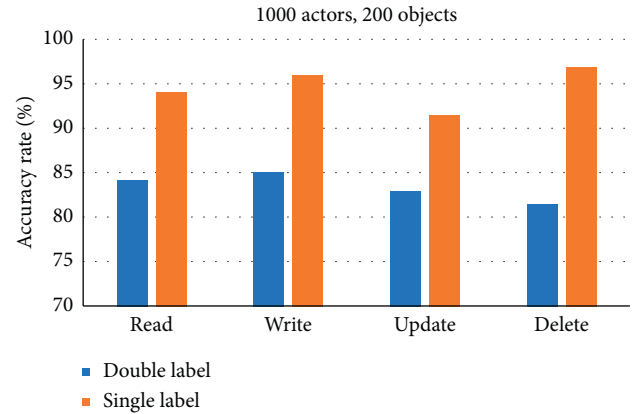


FIGURE 6: Accuracy rates for 1000 actors and 200 objects.

In Table 3, the success of the proposed model (single label) and that of the model in published literature are compared in terms of accuracy. Accuracy rates have been calculated for about 10000 actors and 2000 objects. The success of the proposed model is clearly shown in Figure 7. When the performances of both methods are compared in terms of accuracy rates for all operations performed on objects, the success of the proposed model can be clearly seen.

In Table 4, the success of the proposed model (single label) and that of the model in published literature are compared in terms of accuracy. Accuracy rates have been calculated for about 100000 actors and 20000 objects. The success of the proposed model is clearly shown in 8. When the performances of both methods are compared in terms of accuracy rates for all operations performed on objects, the success of the proposed model is clearly seen.

4.2. Time. In Table 5, the success of the proposed model (single label) and that of the model in published literature in terms of time are compared against the actual data set taken from the hospital. Performances related to time are given for about 100 actors and 20 objects. The success of the proposed model is clearly shown in Figure 9. In terms of time, it is seen that operations are performed on the data in less time with the proposed model. Writing and updating operations take longer in both methods in terms of time compared to other operations. This is because performing writing and reading operations on the object takes more time. Also, when compared in terms of time, the proposed model gives very successful results for all operations performed on the object.

In Table 6, the success of the proposed single-label model and that of the model in published literature in terms of time

TABLE 3: Accuracy rates for 10000 actors and 2000 objects.

Accuracy rate (%)	Double label	Single label
Read	75.63	92.64
Write	81.05	91.09
Update	74.87	89.75
Delete	83.78	90.56

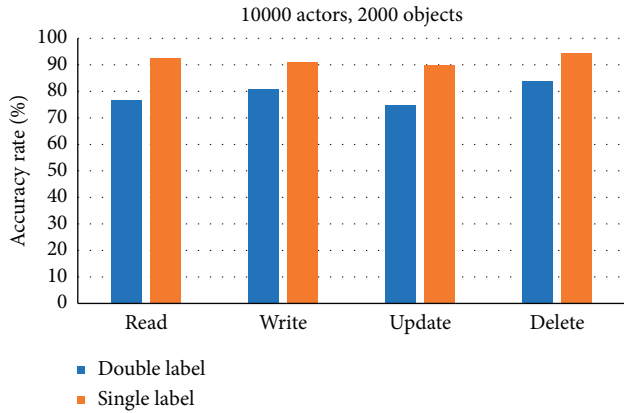


FIGURE 7: Accuracy rates for 10000 actors and 2000 objects.

TABLE 4: Accuracy rates for 100000 actors and 20000 objects.

Accuracy rate (%)	Double label	Single label
Read	70.50	88.64
Write	75.19	87.17
Update	72.65	81.60
Delete	81.64	86.38

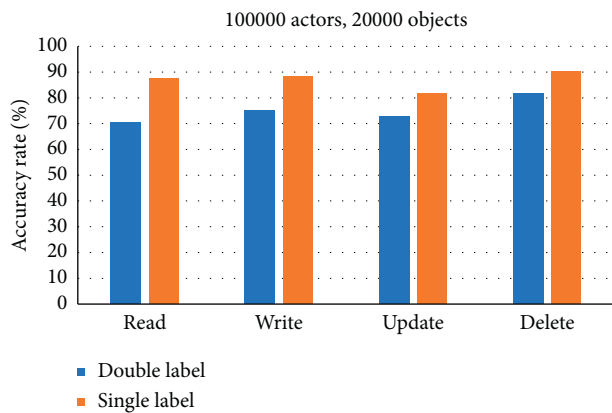


FIGURE 8: Accuracy rates for 10000 actors and 2000 objects.

are compared against the actual data set taken from the hospital. Performances related to the time are given for about 1000 actors and 200 objects. The success of the proposed model is clearly shown in Figure 10. In terms of time, it is seen that operations are performed on the data in less time with the proposed model.

In Table 7, the success of the proposed single-label model and that of the model in published literature in terms of time

TABLE 5: Times for 100 actors and 20 objects.

Time (sec)	Double label	Single label
Read	8.19	6.45
Write	9.82	7.97
Update	12.41	8.60
Delete	6.37	4.84

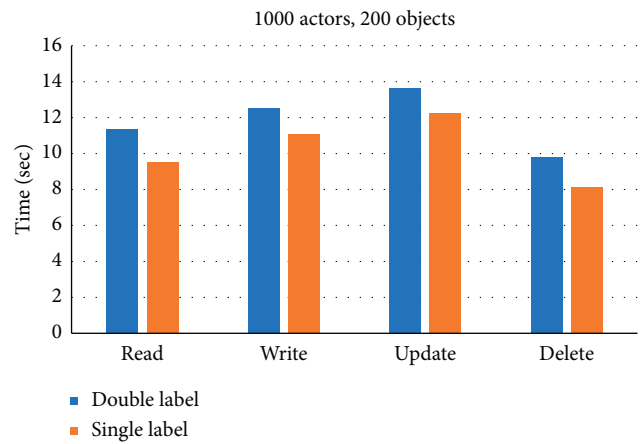


FIGURE 9: Times for 100 actors and 200 objects.

TABLE 6: Times for 1000 actors and 200 objects.

Time (sec)	Double label	Single label
Read	11.35	9.51
Write	12.51	11.09
Update	13.64	12.27
Delete	9.79	8.15

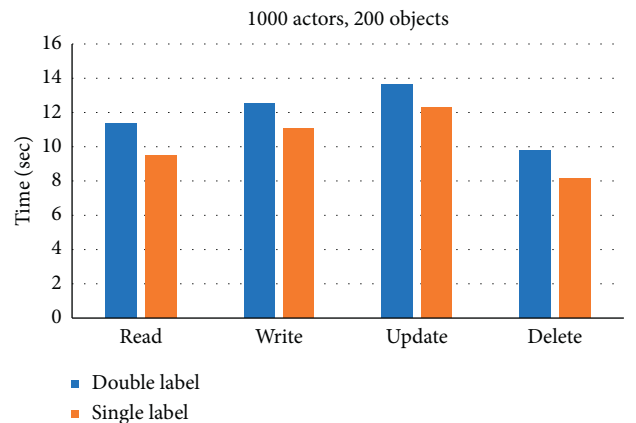


FIGURE 10: Times for 1000 actors and 200 objects.

are compared against the actual data set taken from the hospital. Performances related to the time are given for about 10000 actors and 2000 objects. The success of the proposed model is clearly shown in Figure 11. In terms of

TABLE 7: Times for 10000 actors and 2000 objects.

Time (sec)	Double label	Single label
Read	14.17	11.77
Write	15.02	12.82
Update	16.79	14.73
Delete	11.96	10.66

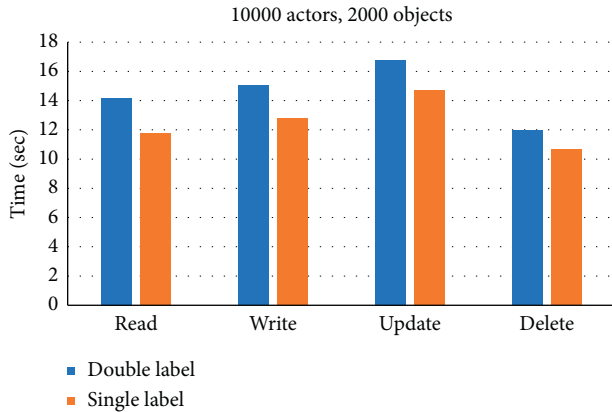


FIGURE 11: Times for 10000 actors and 2000 objects.

TABLE 8: Times for 100000 actors and 20000 objects.

Time (sec)	Double label	Single label
Read	15.72	14.35
Write	17.04	15.37
Update	17.78	16.14
Delete	12.01	11.02

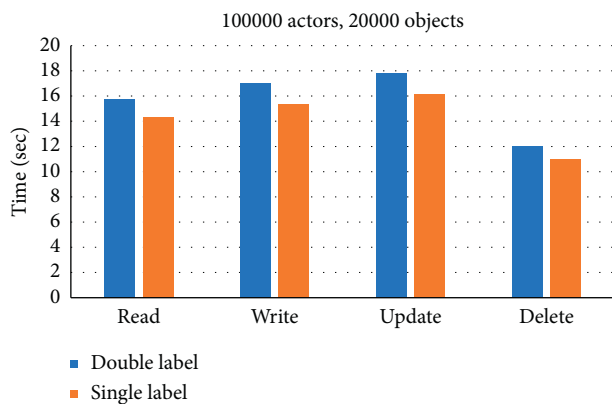


FIGURE 12: Times for 100000 actors and 20000 objects.

time, it is seen that operations are performed on the data in less time with the proposed model.

In Table 8, the success of the proposed single-label model and that of the model in published literature in terms of time are compared against the actual data set taken from the hospital. Performances related to the time are given for

about 100000 actors and 20000 objects. The success of the proposed model is clearly shown in Figure 12. In terms of time, it is seen that operations are performed on the data in less time with the proposed model.

5. Evaluation and Conclusions

In this study, a single-label model was introduced for ensuring data security. In the proposed model, authorization and deauthorization operations between actors were both carried out. Also, in the proposed model, there is no separate authorization or access control for each operation such as reading, writing, updating, and deleting. Access control and authorization operations were performed through labels. Unlike previous studies, data security was ensured for all operations performed in the distributed database. Actors can take back the authority that they give at any time, or they can give authority to the actor they want. Challenges that occur during the implementation of security policies on distributed databases are overcome.

In this study, the problem of data security in distributed databases was addressed. In particular, a distributed-label model related to data flow control was introduced and examples of applications for its use were shown. In addition, data object flows in a distributed environment were modeled with a graph structure. In previous studies, a separate label has been used for each operation (read, write) carried out on the object, and only reading and writing have been performed. In the study proposed here, on the other hand, all operations performed on the object (read, write, update, and delete) were carried out using a single label. This also shows that the proposed model is flexible. By tracking the access of malicious actors to data, attempts were made to prevent disclosure of information. The results of the proposed single-label model for all operations performed on the data were also shown by the experimental study. It delivered more successful results, especially in reading and deleting operations.

The proposed model was also compared with the method used in previous studies in terms of time, and it was seen that it performed operations in a shorter time. In this way, data confidentiality, integrity, and consistency were ensured.

As a future study, a prototype application will be created, which shows the work of the label model, and the model will be enriched by relabeling, which takes into account the hierarchy of actors as well.

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] A. Gupta and E. Galinkin, "Green lighting ML: confidentiality, integrity, and availability of machine learning systems in deployment" in *Proceedings of the 37th International*

- Conference on Machine Learning Workshop on Challenges in Deploying and monitoring Machine Learning Systems*, pp. 1–10, Vienna, Austria, 2020.
- [2] A. Tchernykh, U. Schwiegelsohn, E.-G. Talbi, and M. Babenko, “Towards understanding uncertainty in cloud computing with risks of confidentiality, integrity, and availability,” *Journal of Computational Science*, vol. 36, pp. 1–10, 2019.
 - [3] M. Jouini, L. B. A. Rabai, and A. B. Aissa, “Classification of security threats in information systems,” *Procedia Computer Science*, vol. 32, pp. 489–496, 2014.
 - [4] M. Aminzade, “Confidentiality, integrity and availability – finding a balanced IT framework,” *Network Security*, vol. 2018, no. 2018, pp. 9–11, 2018.
 - [5] O. J. A. Pinno, A. R. A. Grégio, and L. C. E. De Bona, “ControlChain: a new stage on the IoT access control authorization,” *Concurrency and Computation Practice and Experience*, vol. 32, pp. 1–23, 2019.
 - [6] B. A. Tama and K.-H. Rhee, “Data mining techniques in DoS/DDoS attack detection: a literature review,” *Information*, vol. 18, no. 8, pp. 3739–3748, 2015.
 - [7] R. von Solms and J. van Niekerk, “From information security to cyber security,” *Computers & Security*, vol. 38, pp. 97–102, 2013.
 - [8] N. Zeldovich, S. Boyd-Wickizer, and D. Mazieres, “Securing distributed systems with information flow control,” in *Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation (NSDI ’08)*, pp. 293–308, San Francisco, CA, USA, April 2008.
 - [9] M. Krohn, A. Yip, M. Brodsky et al., “Information flow control for standard OS abstractions,” in *Proceedings of ACM Symposium on Operating Systems Principles (SOSP’07)*, vol. 41, pp. 321–334, Stevenson, WA, USA, October 2007.
 - [10] N. Zeldovich, S. Boyd-Wickizer, E. Kohler, and D. Mazières, “Making information flow explicit in HiStar,” *Communications of the ACM*, vol. 54, no. 11, 2011.
 - [11] I. Roy, D. E. Porter, M. D. Bond, K. S. McKinley, and E. Witchel, “Laminar: practical fine-grained decentralized information flow control,” in *Proceedings of the 2009 ACM SIGPLAN Conference on Programming Language Design and Implementation in PLDI’09*, Dublin, Ireland, June 2009.
 - [12] M. Alizadeh, S. Abolfazli, M. Zamani, S. Baharun, and K. Sakurai, “Authentication in mobile cloud computing: a survey,” *Journal of Network and Computer Applications*, vol. 61, pp. 59–80, 2016.
 - [13] M. Rana, M. Kubbo, and M. Jayabalan, “Privacy and security challenges towards cloud based access control in electronic health records,” *Asian Journal of Information Technology*, vol. 16, no. 2–5, pp. 274–281, 2017.
 - [14] Q. Li, R. Sandhu, X. Zhang, and M. Xu, “Mandatory content access control for privacy protection in information centric networks,” *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 5, pp. 494–506, 2017.
 - [15] W. Cheng, R. K. Ports, and D. Schultz, “Abstractions for usable flow control in aelous,” in *Proceedings of the USENIX Conference on Annual Technical Conference (USENIX ATC’12)*, pp. 1–12, Berkeley, CA, USA, June 2012.
 - [16] J. Parker, N. Vazou, and M. Hicks, “LWeb: information flow security for multi-tier web applications,” *Proceedings of the ACM on Programming Languages*, vol. 3, pp. 1–30, 2019.
 - [17] J. Yang, T. Hance, and T. H. Austin, “Armando solar-lezama, cormac flanagan, and stephen chong, precise, dynamic information flow for database-backed applications,” in *Proceedings of the 37th ACM SIGPLAN Conference on Programming Language Design and Implementation*, pp. 631–647, Santa Barbara, CA, USA, June 2016.
 - [18] D. Muthukumar, O. ’K. Dan, C. Priebe, D. Evers, B. Shand, and P. Peter, “FlowWatcher: defending against data disclosure vulnerabilities in web applications,” in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security (CCS ’15)*, Seoul, Republic of Korea, October 2015.
 - [19] E. Cecchetti and A. C. Myers, “Nonmalleable information flow control,” in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS ’17)*, pp. 1875–1891, Dallas, TX, USA, October 2017.
 - [20] J. Liu and O. Arden, “Fabric: building open distributed systems securely by construction,” *Journal of Computer Security*, vol. 25, no. 4-5, pp. 367–426, 2017.
 - [21] A. C. Myers and B. Liskov, “Protecting privacy using the decentralized label model,” *ACM Transactions on Software Engineering and Methodology*, vol. 9, no. 4, pp. 410–442, 2000.
 - [22] D. Servos and S. L. Osborn, “Current research and open problems in attribute-based access control,” *ACM Computing Surveys*, vol. 49, no. 4, 2017.
 - [23] J. Liu and M. D. George, “Fabric: a platform for secure distributed computation and storage,” in *Proceedings of the ACM Symposium on Operating Systems Principles and Implementation (SOSP)*, pp. 321–334, Koblenz, Germany, October 2009.
 - [24] N. Burow, S. A. Carr, J. Nash et al., “Control-flow integrity,” *ACM Computing Surveys*, vol. 50, no. 1, pp. 1–33, 2017.
 - [25] Q. Aafaf, M. Hajar, A. E. Anas, and A. Q. Abdellah, “Access control in the internet of things: big challenges and new opportunities,” *Elsevier Computer Networks*, vol. 12, pp. 237–262, 2017.
 - [26] J. B. D. Joshi, E. Latif, and A. Ghafoor, “A generalized temporal role-based access control model,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 1, pp. 4–23, 2005.
 - [27] M. R. Anawar, S. Wang, M. Azam Zia, A. K. Jadoon, U. Akram, and S. Raza, “Fog computing: an overview of big IoT data analytics,” *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 7157192, 22 pages, 2018.
 - [28] M. A. Albahar, “Recurrent neural network model based on a new regularization technique for real-time intrusion detection in SDN environments,” *Security and Communication Networks*, vol. 2019, Article ID 8939041, 9 pages, 2019.
 - [29] L. Rodrigues Prete and C. M. Schweitzer, “Simulation in an SDN network scenario using the POX Controller,” in *Proceedings of the 2014 IEEE Colombian Conference on Communications and Computing (COLCOM)*, Bogota, Colombia, June 2014.
 - [30] A. C. Myers and B. Liskov, “Complete, safe information flow with decentralized labels,” in *Proceedings of the 1998 IEEE Symposium on Security and Privacy*, pp. 1–12, Oakland, CA, USA, May 1998.
 - [31] O. Arden and A. C. Myers, “A calculus for flow-limited authorization,” in *Proceedings of the 2016 IEEE 29th Computer Security Foundations Symposium (CSF)*, pp. 135–149, Lisbon, Portugal, June 2016.
 - [32] A. C. Myers and B. Liskov, “A decentralized model for information flow control,” *ACM SIGOPS Operating Systems Review*, vol. 31, no. 5, pp. 129–142, 1997.